

분자진화기술을 이용한 단백질 개량

김의중, 정흥채, 반재구 / (주)제노포커스

Genome 시대의 기술발전과 방대한 자연계의 biodiversity로부터 신규 유용 유전자를 발굴하는 기술은 비약적으로 발전하고 있다. 초고속 발현, 기능 분석, 단백질간의 상호작용 연구, 데이터베이스를 구축하는 연구도 중요하지만, 찾은 유전자를 산업적으로 응용할 수 있도록 빠른 시간 내에 최적화시키는 high-throughput biology가 매우 중요하게 되었다. 그 핵심 기술은 분자진화 (directed evolution) 기술이다.

Directed evolution 기술은 타겟 단백질 또는 생명체를 우리가 원하는 특성을 갖도록 초고속으로 개량하는 기술이다. 자연계에서 발굴한 유전자를 실제 목적에 바로 이용하기에는 여러 가지 제약이 따르는 경우가 많다. 예를 들어, 분리한 효소가 그 효소의 이용분야에 비활성, 기질 특이성, pH 및 열 안정성 등이 최적인 경우가 드물다. 이런 경우, 과거에는 탐색한 효소에 적합하도록 공정을 개선하거나 처음부터 새로운 효소를 스크리닝하는 경우가 많았다. 하지만 방대한 biodiversity로부터 신규 효소 유전자원을 쉽게 탐색할 수 있고, 이들 유전자원으로부터 원하는 활성을 갖도록 directed evolution이 가능하게 됨으로써 'ideal' 한 공정에 맞는 이상적인 효소를 디자인할 수 있도록 'paradigm shift' 가 일어나고 있다 [그림 1].

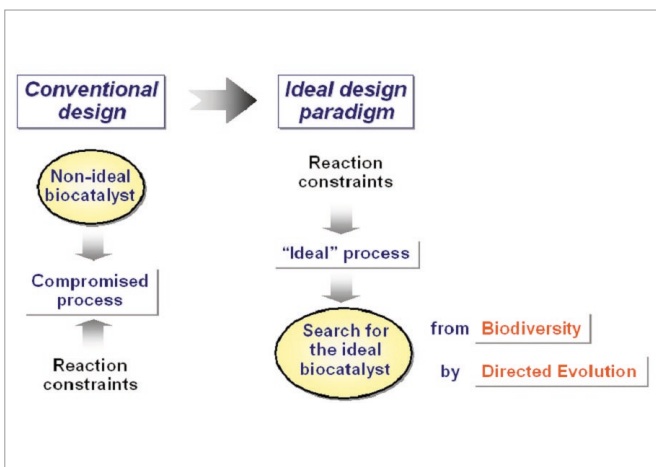


그림 1. 이상적인 효소반응 구축을 위한 paradigm shift.

Directed evolution 개념은 80년대 이미 소개되었으며, PCR을 행할 때 error가 유발되도록 하는 error-prone PCR을 이용한 단백질 개량 시도가 있었다. 생명공학분야에 핵심기술로서 directed evolution이 각광받게 된 것은 1994년 Stemmer에 의해서 DNA shuffling 기술이 개발된 이후부터라고 해도 과언이 아니다. DNA shuffling은 다수의 유전자원으로부터 장점을 조합하여 combinatorial chimera를 제조할 수 있기 때문에, evolution 속도를 혁신적으로 높일 수 있게 되었다. 따라서, 포스트제분 시대의 방대한 유전자원으로부터 원하는 특성을 갖는 최적 유전자를 얻기 위해서 directed evolution 기술이 필수적으로 이용되는 시대가 된 것이다.

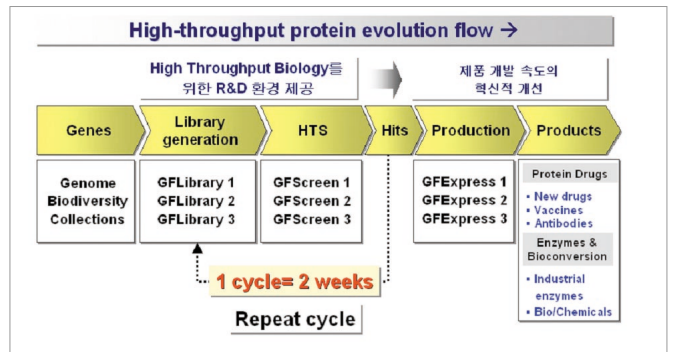


그림 2. Directed evolution을 이용한 제품개발 흐름도.

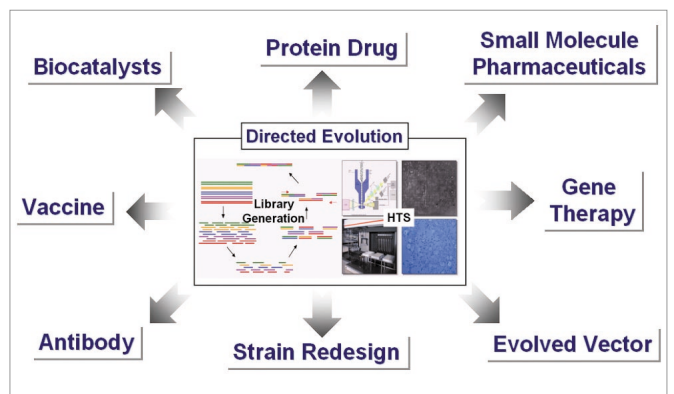


그림 3. Directed evolution의 응용분야.

Directed Evolution 기술의 핵심

Directed evolution을 기술적으로 다시 설명하면, i) 유전자원에 각종 돌연변이를 주어 라이브러리를 만든 후, ii) 이 라이브러리내에 우리가 원하는 활성으로 개량된 변이체를 초고속으로 스크리닝하는 두 단계를 반복함으로써, 최종적으로 우리가 원하는 최적 활성의 단백질을 얻게 되는 기술이다 [그림 4].

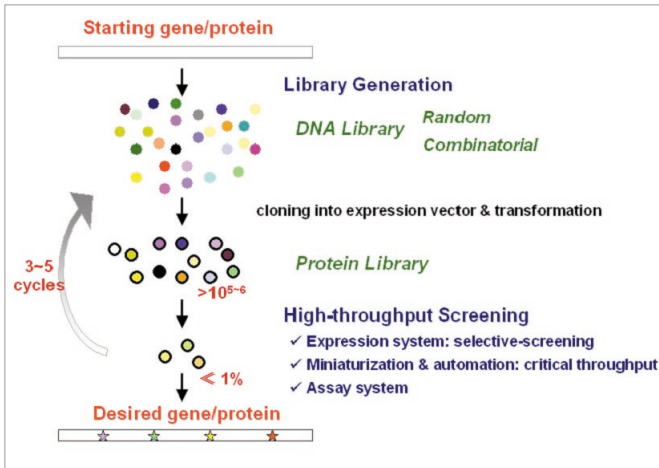


그림 4. Library generation과 screening흐름도

본 고에서는 directed evolution의 핵심기술인 라이브러리 제조기술과 초고속 스크리닝기술의 발전을 리뷰함으로써 성공적인 단백질, 특히 효소의 연구, 개발, 산업화를 위한 directed evolution의 중요성을 강조하고자 한다.

라이브러리 제조

단백질 개량 속도는 1차적으로 단백질 유전자에 돌연변이를 가하여 다양한 변이체, 즉 라이브러리를 제조하는 방법에 따라 결정되며, 이는 곧 원하는 활성을 지니는 단백질을 얼마나 빠르게 얻을 수 있는가와 직결된다. 라이브러리를 제조하는 방법은 크게 유전자에 무작위적인 돌연변이를 일으키는 무작위돌연변이법 (random mutagenesis)과 서로 다른 2가지 이상의 변이체간의 조합을 통하여 새로운 돌연변이를 제조하는 조합돌연변이법 (combinatorial mutagenesis)이 있다 [그림 5].

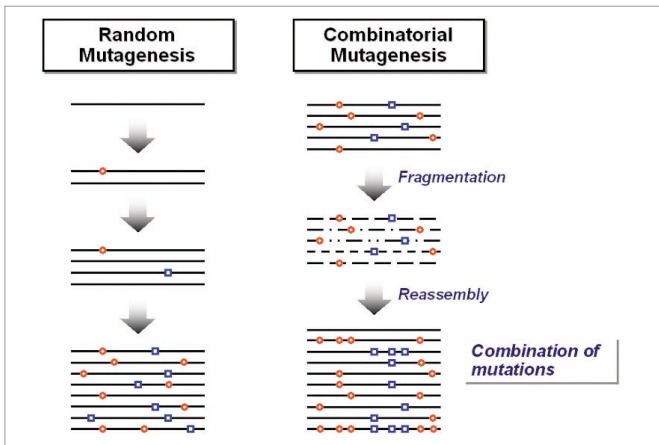


그림 5. 무작위돌연변이법과 조합돌연변이법.

무작위돌연변이법은 i) EMS, NTG 등 돌연변이 유발물질 (mutagen)을 사용하는 화학적 방법, ii) 유전자증폭 반응중에 변이의 확률을 증가시키는 error-prone PCR 법, iii) DNA 복제시 repair에 관련된 유전자가 결손된 숙주세포를 이용하는 mutator strain법 등이 있다. 이 중 가장 널리 사용되는 것은 error-prone PCR법으로서, PCR을 행할 때 Mn^{2+} 이온의 농도를 높이고, dNTP의 불균형적인 조합을 통하여 Taq 중합효소의 본연의 에러발생율보다 높은 돌연변이가 생성되도록 하는 것이다 (1~8 /1 kb).

이러한 무작위돌연변이법으로 제조된 라이브러리에서 스크리닝한 최고 활성의 돌연변이체는 일반적으로 내가 원하는 활성에 도움이 되는 긍정적인 돌연변이 (positive mutation) 외에도 해가 되는 부정적 돌연변이 (negative mutation)가 공존하기 때문에, 라이브러리 제조 및 스크리닝을 반복함에 따라 부정적 돌연변이의 축적으로 단백질 개량 속도가 느려진다. 반면에 조합돌연변이법은 라이브러리 제조시 긍정적 돌연변이의 조합뿐만 아니라, 부정적 돌연변이의 제거로 인하여 보다 높은 활성의 변이체를 쉽게 개발할 수 있도록 한다 [그림 6].

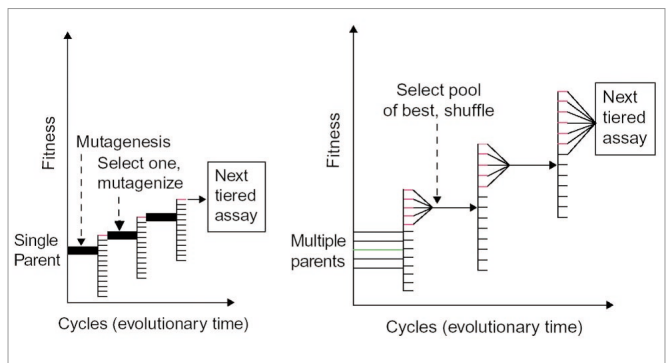


그림 6. 무작위돌연변이법과 조합돌연변이법을 이용한 단백질 분자진화 속도.

조합돌연변이법은 1994년 Stemmer에 의해서 처음 개발되었다. 이 기술은 염기서열이 비슷한 유전자들을 DNase로 절단하여 다양한 DNA 단편들을 제조하고, 이 단편들을 이용하여 별도의 primer 첨가 없이 PCR을 하게 되는데, 이때 각각의 단편들은 유사한 염기서열을 가지고 있는 다른 단편에 대하여 각기 primer로 작용하게 된다. 이렇게 재조합된 DNA 단편들은 다시 말단에 결합하는 primer를 이용한 PCR을 통하여 전체길이를 갖는 새로운 유전자로 조합되게 되는데, 이러한 일련의 과정이 마치 카드를 섞듯이 유전자 조각들을 무작위로 섞어서 매우 다양한 조합 라이브러리를 구축한다고 하여 DNA shuffling이라고 명명되었다 [그림 7]

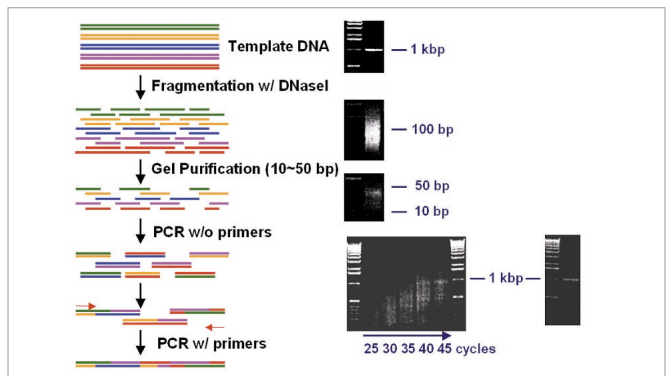


그림 7. DNA shuffling 기술의 모식도.

DNA shuffling 기술이 개발된 이후, 어떻게 주형 DNA를 절단할 것인가 (DNase 사용, PCR 이용, 제한효소 사용, 물리적 분해 등), 그리고 어떻게 다시 조합시킬 것인가 (PCR, ligation, recombination 등)에 따라서 StEP, Random Priming, RCR, CLERY, L-Shuffling, RETT, GFLibrary 등 매우 다양한 조합돌연변이법들이 개발되었다. DNA shuffling은 염기서열이 70~80% 이상의 상동성을 가지는 DNA간에만 조합이 가능한데, 전혀 상동성이 없는 두개의 DNA간에 조합돌연변이를 일으킬 수 있는 ITCHY [그림 8], SHIPREC, RACHITT 기술 등도 개발되어 이용되고 있다.

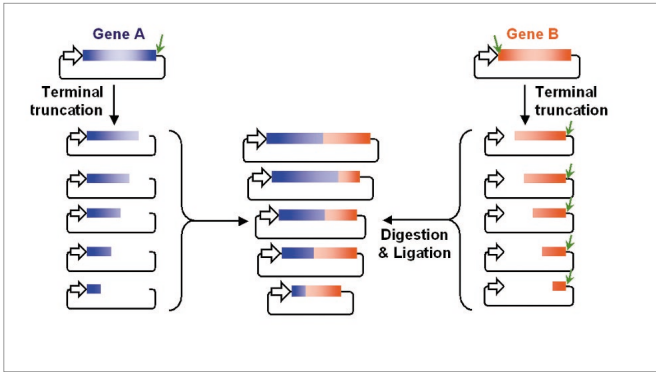


그림 8. ITCHY 기술의 모식도.

최근 라이브러리 제조 기술은 가능한 적은 규모의 라이브러리에 활성이 개량된 돌연변이를 보다 많이 포함하도록 제조하는 것이 목표가 되고 있다. 이는 무작위돌연변이법과 조합돌연변이법의 장단점을 적절히 조화시키고, 단백질 구조 및 활성에 대한 데이터베이스에 기반한 바이오인포매틱스를 적용하여 최적의 라이브러리를 컴퓨터로 디자인함으로써 가능해지고 있다.

스크리닝

왜 스크리닝 기술이 중요해지고 있는가? ① metagenome 등의 생물다양성 확보와 계통기술의 진보, ② 실험실적 분자다양성 생성기술의 진보, 그리고 ③ 컴퓨터의 발달로 디자인 가능한 분자다양성 제조기술 발달로 라이브러리 제조능력이 향상됨에 따라 혁신적인 스크리닝 방법이 필요하다. 단백질 개량을 위하여 무엇을 스크리닝하는가? 스크리닝의 기준은 효소의 역가 (activity), 안정성 (stability), 생산성 (availability) 등이 증가한 변이체를 탐색한다.

일반적인 단백질 개량을 위한 스크리닝과정은 ① 타겟유전자의 확보, ② 유전자에 무작위 변이 도입, ③ 적당한 숙주세포에 라이브러리를 형질전환하여 발현하기, ④ 개선된 효소활성을 스크리닝하기, ⑤ 탐색된 효소의 활성 확인하기, ⑥ 추가개선을 위한 2차 변이 도입 및 재탐색하기 등으로 구성되어 있다.

얼마나 많은 라이브러리를 스크리닝할 것인가? 일반적으로 유전자에 변이를 가하면 효소활성이 살아남는 비율은 1~10% 정도이고 그 중에 활성이 증가한 경우는 0.1~1% 정도에 지나지 않아 실제로는 총 라이브러리 중 0.001~0.1% 정도에 지나지 않는다. 따라서 10개의 positive clone을 확보하기 위해서는 만개에서 백만개의 라이브러리를 스크리닝해야 한다 [그림 9]. 하지만 개량해야 하는 유전자, 라이브러리 제조방법, 개량 포인트에 따라 달라진다는 점을 간과해서는 안된다.

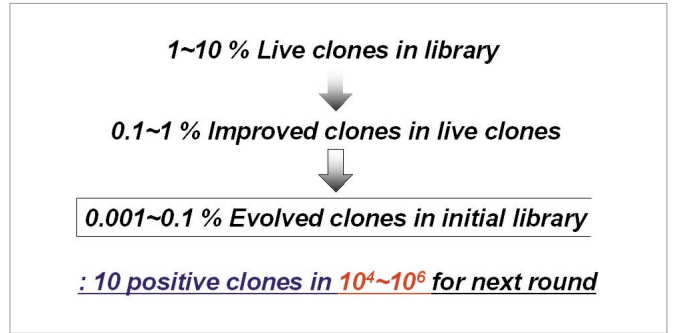


그림 9. 얼마나 많은 라이브러리를 스크리닝해야 하나?

스크리닝기술에 따라 어느 정도의 라이브러리 처리속도 (throughput)에 도달할 수 있는가? 대량의 라이브러리로부터 타겟 클론의 스크리닝은 처리속도가 낮은 전통적인 분석기로부터 flow cytometer와 같은 고속 대용량으로 처리할 수 있는 기기까지 발전하였다 [표 1 참조].

Number of individual candidate clones	Assay Method
>10 ⁷	Genetic selection, FACS, Display
10 ⁴⁻⁶	Solid state assays: colorimetric or fluorescence
10 ²⁻⁴	Microtiter format assays: colorimetric or fluorescence
1-10 ²	Individual high precision assays: GC, HPLC, mass spectrometry

[J. Minshull and W.P.C. Stemmer, Current Opinion in Chemical Biology 1999, 3:284-290.]

표 1. 스크리닝방법에 따른 처리속도비교

Microplate assay

가장 대표적인 고속 스크리닝법으로 96 well plate를 가장 많이 활용하고 있으나 점차 386 well plate의 활용이 증가되고 있다. 예를 들어 NAD(P)H와 연계한 oxidoreductase 라이브러리의 초고속 스크리닝에 발색반응계를 활용하면 96 well plate에서 쉽게 라이브러리를 스크리닝할 수 있다 [그림 10].

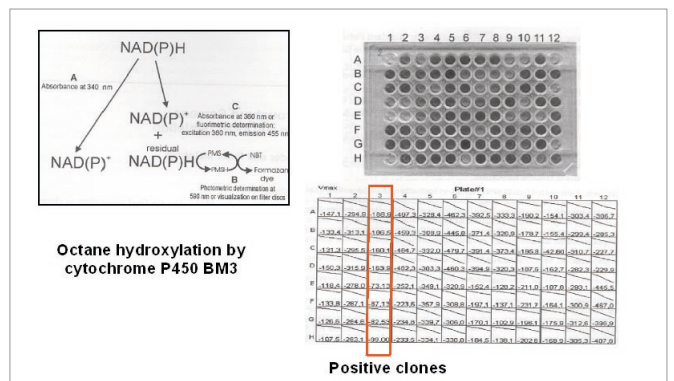


그림 10. Microtiter format screening.

미세로봇과 같은 특수한 경우는 수천에서 십만 well이 있는 플레이트를 사용하여 처리속도를 높이고 있다 [그림 11].

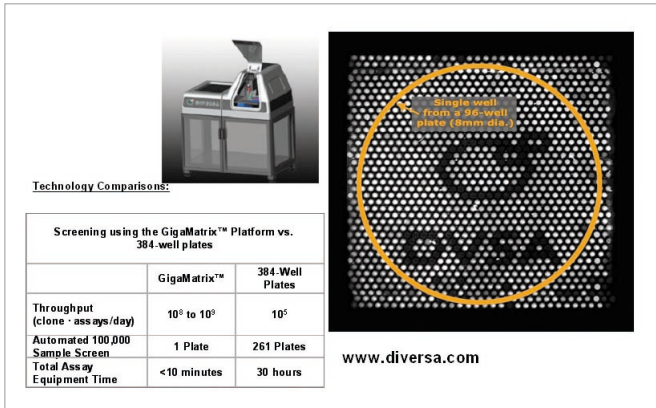


그림 11. Diversa의 GigaMatrix™ 플레이트.

Solid state screening

한국생명공학연구원 최의성 박사팀은 리파제의 라이브러리를 효모의 표면에 발현하여 직접 플레이트 고체배지에서 halo를 형성하는 정도를 효소활성과 연계하여 스크리닝할 수 있는 방법을 활용하여 리파제의 비활성이 증가한 변이체를 쉽게 확보할 수 있었다 [그림 12].

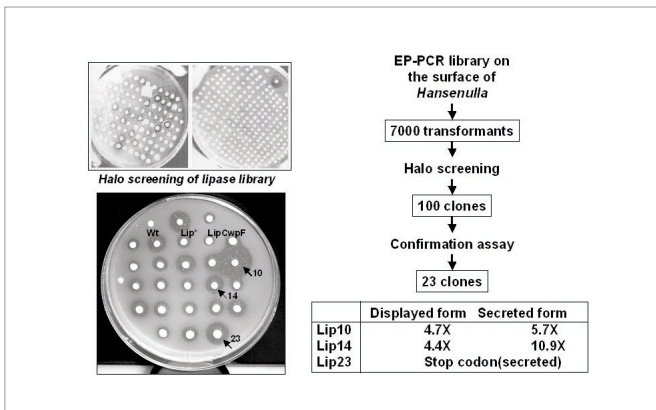


그림 12. Solid state screening: Halo 형성을 통한 리파제 라이브러리 스크리닝.

Kairos사 (미국)는 Macroscopic digital imaging spectrophotometer를 활용하여 고체 평판배지에 형성된 콜로니를 특정 파장에서 스캔함으로써 흡광정도를 이미지로 전환 스크리닝하는 기술을 상업화하였다 [그림 13].

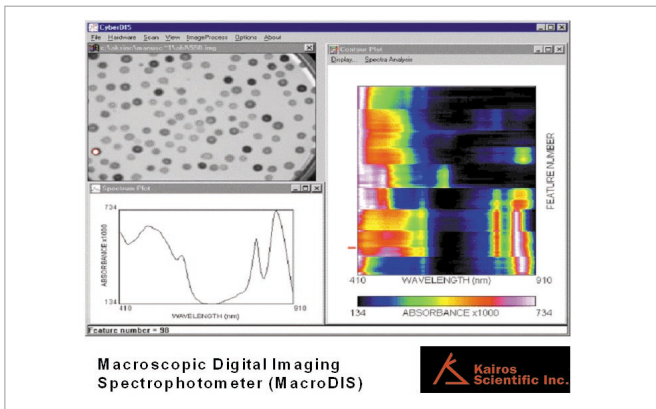


그림 13. Kairos사의 digital imaging spectrophotometer를 이용한 스크리닝기술.

Genetic selection

미생물의 죽고 사는 문제와 연결할 수 있는 장점이 있는 genetic selection은 전통적으로 항생제 마커나 미생물의 대시중에 성장과 직접연계한 효소가 knockout된 변이주를 이용한 경우가 많았다. 최근 chloramphenicol (Cm)에 저항성을 보이는 CAT 유전자를 타겟 단백질과 fusion하여 발현하고 Cm 배지에서 살아남는 것을 스크리닝하는 방법을 개발하여, 인간 유래 단백질의 solubility가 증가한 변이체를 스크리닝하는 기술로 활용되고 있다 [그림 14].

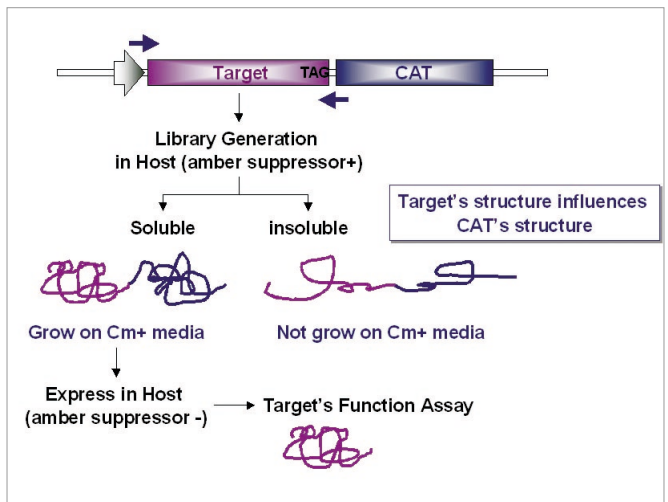


그림 14. CAT-fusion을 이용한 단백질의 solubility 스크리닝 기술.

Display-based screening

단백질을 생체 표면에 부착하여 발현시킬 수 있는 표면발현기술 (display technology)은 세포표면에 발현된 단백질이 외부에서 넣어준 기질이나 리간드와 결합이 용이하기 때문에 단백질의 초고속 스크리닝 툴로서 광범위하게 사용되고 있다. 특히 항체나 효소 라이브러리를 디스플레이하여 결합력, 활성, 안정성이 증가한 변이체를 biopanning법, flow cytometer를 이용한 sorting법 등으로 초고속으로 스크리닝해내고 있다 [그림 15, 16]

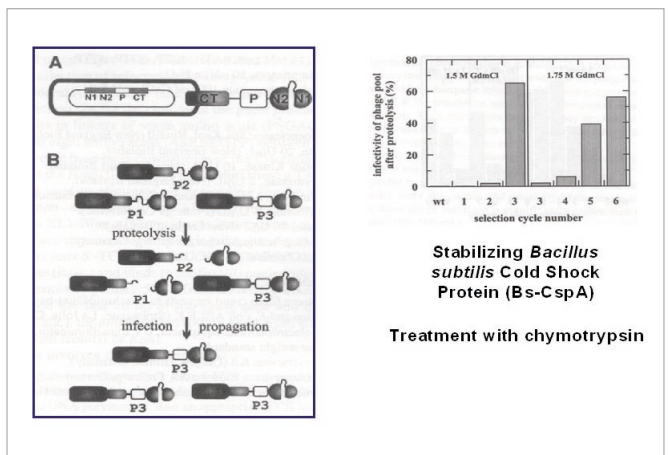


그림 15. Biopanning법을 이용한 단백질의 안정성 증가 스크리닝.

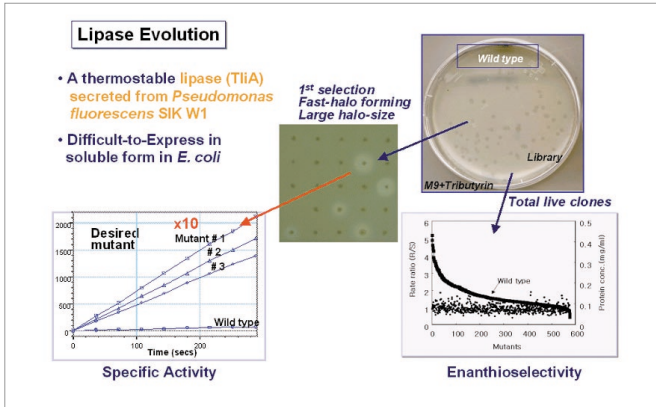
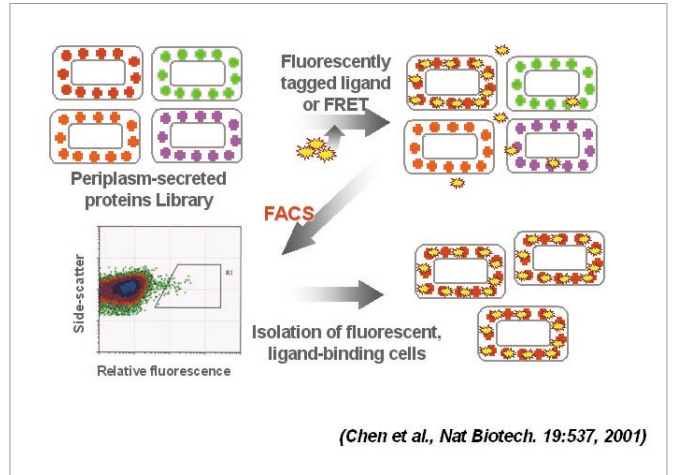


그림 16. 박테리아 표면발현기술을 이용한 리파제의 비활성 및 enantioselectivity 증가 스크리닝.



(Chen et al., Nat Biotech. 19:537, 2001)

그림 19. Periplasm 발현과 FACS를 이용한 초고속 스크리닝 기술.

HTS instruments

최근에는 다양한 초고속 스크리닝이 가능한 정밀기기의 발달로 하루 수십~수백만개 이상의 라이브러리를 처리할 수 있게 되었다. 처리속도를 증가시키기 위한 기본적인 기기들은 그림 17에 보인 것처럼 기존의 손에 의존한 과정을 대신하는 로봇들이다. 특히 스크리닝에 필수적인 미생물 배양 단계를 대량으로 할 수 있는 High-throughput culture system인 MegaGrow와 같은 기기는 주목할 만하다.

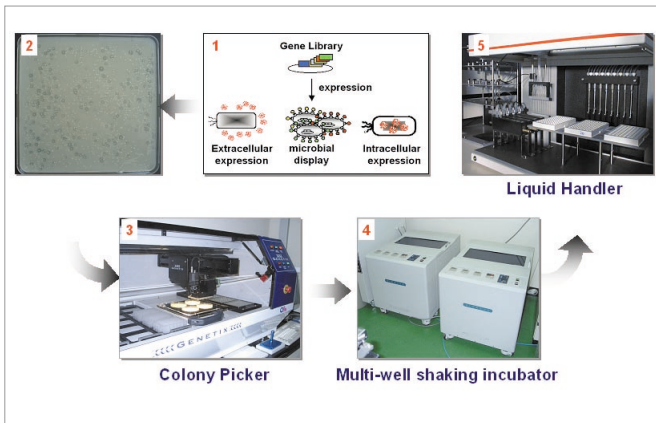


그림 17. 스크리닝 처리속도를 높이기 위한 특수발현시스템과 1) 기기들, 2) Large square plate, 3) Colony picking robot, 4) High-throughput culture system, 5) Liquid handling robot.

Flow cytometer를 이용한 라이브러리의 sorting은 현재 시간당 10⁶개의 속도로 스크리닝할 수 있는 환경을 제공할 수 있다 [그림. 18]

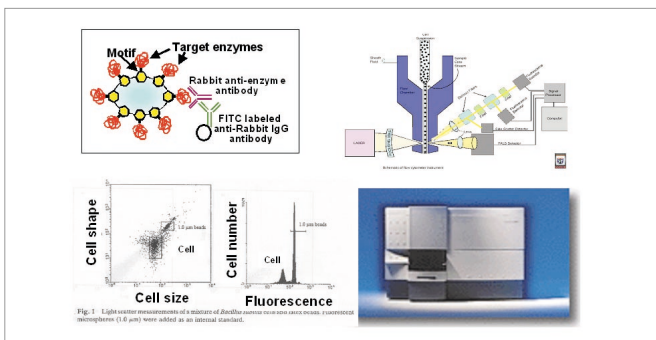


그림 18. Flow cytometer, fluorescence-activated cell sorter (FACS).

특히 대장균의 periplasm에 효소나 항체의 라이브러리를 발현하고 FACS를 이용하여 초고속으로 활성이나 결합력이 증가한 변이체를 스크리닝할 수 있음을 보였다 [그림 19].

반응액의 부피를 nanoliter까지 줄여 대량의 라이브러리를 스크리닝할 수 있는 chip의 개발은 스크리닝 처리속도를 획기적으로 증가시키고 있다 [그림 20, 21].

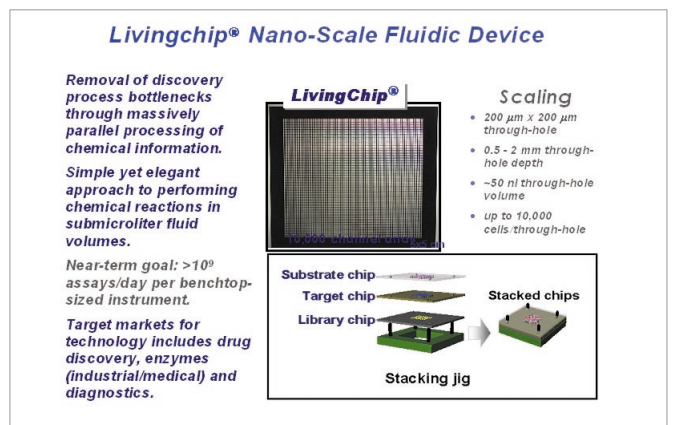


그림 20. Nanoliter scale fluidic device, Livingchip™ (BioTrove사, 미국).

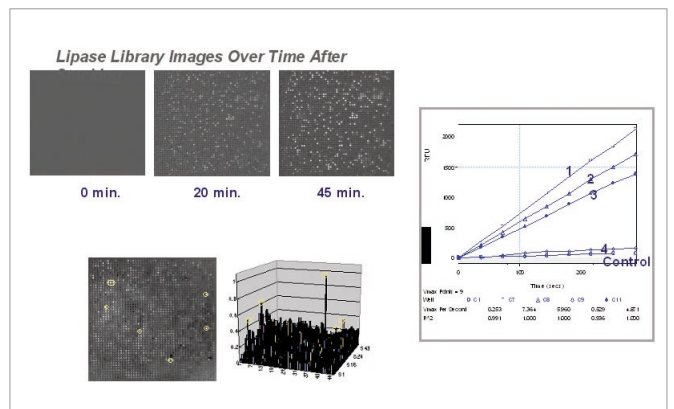


그림 21. Livingchip을 이용한 대장균에 디스플레이된 리파제 라이브러리의 스크리닝에 (리파제의 비활성 증가 변이체 스크리닝).

White biotechnology 시대의 directed evolution

Directed evolution 기술은 날로 높아지는 biotransformation 기술 수요와 맞물려서 더욱 각광을 받고 있다. Biotransformation 기술은 i) green chemical 및 natural chemical 합성 분야, ii) 화학으로 불가능한 반응공정의 실현, iii) 경제성이 확보되는 새로운 공정 개발 등에 적용이 가능하다. 이제는 Directed evolution 기술의 발전으로 원하는 활성을 갖는 효소를 보다 쉽게 찾거나, 개량하거나, 잘 '디자인' 할 수 있게 되었고, 이는 화학 산업에서의 바이오기술 응용을 촉진하고 있다. 이러한 'White Biotechnology' 분야는 공해 저감이나 재생가능자원활용이라는 관점에서 뿐만 아니라 일상적인 산업제품의 부가가치를 높이는 데 biotransformation 기술의 활용가능성 또한 커질 것으로 기대된다. 이 모든 기술 발전의 핵심은 directed evolution을 통한 효소디자인과 효소를 잘 활용할 수 있는 biotransformation 반응을 디자인하는 것이다. 결론적으로 성공적인 biotransformation 공정을 개발하는데 가장 중요한 것은 공급가능한 최적의 효소를 제조하는 것이며, 이를 위한 핵심기술로서 directed evolution을 이용한 맞춤형 효소 개발이 필수적으로 필요하게 되었다 [그림 22].

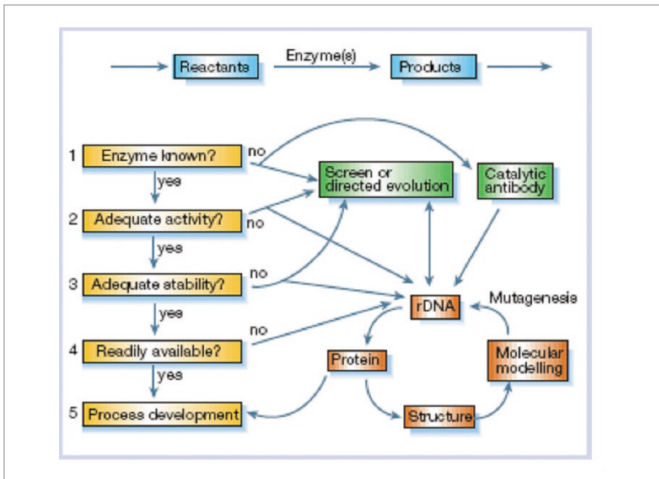


그림 22, Biotransformation 효소개발 전략.

참고문헌

- Cherry J.R. and Fidantsef A.L. 2003. Directed evolution of industrial enzymes: an update. *Curr. Opin. Biotechnol.* **14**:438-43.
- Ball P. et al., 2001. Biocatalysis: Synthesis methods that exploit enzymatic activities. *Nature* **409**:225-268
- Olsen M.J. et al. 2000. High-throughput screening of enzyme libraries. *Curr. Opin. Biotechnol.* **11**:331-337.
- Daugherty, et al. 2000. Flow cytometric screening of cell-based libraries. *J. Immunol. Methods* **243**:211-227.
- Cipolla, L. 2004. Combinatorial libraries of biocatalysts: application and screening. *Comb. Chem. High. Throughput. Screen.* **7**:101-114.
- Dalby P.A. 2003. Optimising enzyme function by directed evolution. *Curr. Opin. Struct. Biol.* **13**:500-505.
- Entzeroth M. 2003. Emerging trends in high-throughput screening. *Curr. Opin. Pharmacol.* **3**:522-529.
- Petrounia I.P. and Arnold F.H. 2000. Designed evolution of enzymatic properties. *Curr. Opin. Biotechnol.* **11**:325-330.

- Arnold F.H. 2001. Combinatorial and computational challenges for biocatalyst design. *Nature* **409**: 253-257.
- Schmidt-Dannert C. and Arnold F.H. 1999. Directed evolution of industrial enzymes. *Trends Biotechnol.* **17**:135-136
- Zhang Y.Z. et al. 2002. Genome shuffling leads to rapid phenotypic improvement in bacteria. *Nature* **415**: 644-646.
- Cramer, A. et al. 1998. DNA shuffling of a family of genes from diverse species accelerates directed evolution. *Nature* **391**:288-291.
- Stemmer, W.P.C. 1994. Rapid evolution of a protein in vitro by DNA shuffling. *Nature* **370**:389-399.



김의중 (ejkim@genofocus.com)

(주)제노포커스 기반기술팀장

1997.3. ~ 2001. 2.
2001.3. ~ 현재

연세대학교 생명공학과 박사
(주)제노포커스 선임연구원



정흥채 (hcjung@genofocus.com)

한국생명공학연구원 선임연구원

1994.3. ~ 1998. 2.
1991.3. ~ 현재
2000.4. ~ 현재

프랑스 공피엔느공대 화학공학과 박사
한국생명공학연구원 선임연구원
(주)제노포커스 이사



반재구 (jgpan@genofocus.com)

(주)제노포커스 대표이사

1982.3. ~ 1985. 8.
1982.8. ~ 1983.12.
1986.3. ~ 1986. 7.
1985.9. ~ 1994. 2.
1992.9. ~ 1992.12.
1994.9. ~ 1995. 2.
1994.9. ~ 2001. 7.
2000.4. ~ 현재

한국과학기술원 생물공학과 박사
프랑스 공피엔느공대 학생연구원
프랑스 공피엔느공대 post-doc
한국생명공학연구원 선임연구원
한국과학기술원 대우교수
한국과학기술원 겸임교수
한국생명공학연구원 책임연구원
(주)제노포커스 대표이사