

생물정보학 기반 miRNA target 유전자 예측

충북대학교 | 류제운 · 유재수* · 김학용

1. 서론

1.1 생물정보학이란?

생물정보학 또는 생명정보학은 매우 다양한 응용분야를 담고 있는 폭넓은 학문이다. 한 문장으로 간략하게 표현하기가 쉽지는 않지만 “생명현상 연구에 필요한 전산학/통계학/수학을 이용하는 것” 또는 “컴퓨터와 소프트웨어를 활용해 생명현상을 연구하는 것”으로 정의할 수 있다[1]. 생물정보학은 다양한 분야의 지식을 필요로 하는 융합학문 중 하나이다. 복잡계의 하나인 생명현상을 이해하는 것은 매우 복잡하고 어려운 일이기 때문에 융합학문 또는 융합기술이 필요하다. 이러한 이유로 등장한 생물정보학은 컴퓨터의 발달과 더불어 분자생물학, 유전공학, 세포생물학 분야 등의 생명과학의 급격한 발전과 게놈 프로젝트의 결과로 엄청난 양의 생물학적 데이터베이스가 쏟아져 나왔기 때문에 다른 융합학문에 비해 훨씬 앞서 융합과학의 면모를 갖추게 되었다. 체계적인 데이터베이스를 구축하고 효율적으로 분석하려는 노력과 더불어 생명과학 분야의 폭넓은 발전과 어깨를 같이하면서 다양한 생명체가 지닌 생명현상의 비밀이 속속 밝혀지고 있다.

1.2 생물정보학의 연구동향

이러한 생물정보학이 구체적으로 어떠한 연구에 주로 쓰이는가에 대해서는 연구자들의 연구 분야와 접근방향에 따라 다양하게 카테고리를 나눌 수 있는데, 그 중에서 가장 널리 알려진 Jones(2000)이 분류한 네 가지를 살펴보면 생물정보학의 연구방향을 살펴보고자 한다[2].

- **서열 데이터로부터 유전자 예측 및 단백질 기능 예측**
염기서열분석기 등을 통해 얻은 데이터를 생물정보학 솔루션 및 DB를 통해 가공하여 유전자 위

치, 서열 정보, 유전자 정보, 아미노산 서열로 번역하여 유전자 DB를 구축한다. 이 과정에서 얻은 정보를 제약 회사 등에게 제공하여 보다 나은 신약 타겟을 선정할 수 있도록 해주며, 식품 회사 등에는 원하는 특정 유전자 삽입 및 형질 전환 등에 대한 정보를 제공함으로써 저비용 고품질 대량생산 시스템의 기회를 제공한다.

- **약물표적 추적**

유전자와 단백질에 대한 연구가 더 진전하고 관련 데이터양의 증가 및 시뮬레이션 툴의 기능의 향상은 약물의 대상이 되는 화합물이 인간의 단백질과 어떻게 결합할 것인가 등을 추론하여 임상 실험 이전의 단계에서 가장 적절한 화합물을 얻을 수 있다[3].

- **유전자·단백질 발현 정보 분석**

유전자와 단백질 발현 정보는 단순한 정성적 분석만이 아닌 정량적 분석, 대사회로 전체의 흐름 등이 필요한 매우 복잡하고 유기적으로 얽혀 방대한 양을 필요로 하는 정보이다. 질병의 이해 및 미생물·가축·식물 등의 대사회로에 있어서 원하는 대사산물의 발현 양과 질을 조절을 위해서 반드시 필요하다. 현재 마이크로어레이(microarray), NGS(Next-generation sequencing)와 같은 high throughput 기술의 발전으로 얻은 대량 발현정보를 데이터베이스화하고 분석하는 알고리즘에 관한 연구가 진행되고 있다[4].

- **개인의 유전자 다양성**

개인 간 유전적 차이를 나타내는 SNP(single nucleotide polymorphism) 정보는 인구 수 만큼이나 방대하고 개인과 인종만큼이나 다양하기 때문에 당장은 상업적으로는 이용이 어려운 측면이 있지만 제약 회사의 신약 개발 실패의 가장 큰 문제점이 부작용에 있다는 점을 감안한다면, 향후 이 분야

* 종신회원

† 본 연구는 KISTI의 지원(K-13-L01-C02)과 질병관리본부 학술연구용역과제(2013E6200100) 연구비를 지원받아 수행되었음.

에 대한 생명과학 및 생물정보학 연구는 꾸준히 증가하리라 예상된다.

1.3 21세기 새로운 생명과학의 화두, 후성유전학

앞에서 살펴본 바와 같이 20세기 생명과학은 DNA 나 단백질 중심으로 연구되었으며 이 결과가 정보과학 과 접맥되면서 급속히 발전되었다. 이제 21세기 생명 과학은 정보과학과 더불어 복잡계 과학으로 전환되고 있는 시점이다. 이에 생명체 자체를 벗어나 생명뿐만 아니라 주위 환경을 고려하는 생명과학에 관한 연구가 새로이 대두되고 있다. 생명체 자체는 DNA 염기서열 의 변화 없이도 환경의 변화에 다양하게 반응하고 적 응하고 있는 사실에 근거를 두고 새로운 시각으로 생 명과학을 연구하는 방향으로 전환되고 있다. 다시 말 해, 게놈 프로젝트가 밝혀낸 선천적인 유전체 정보만이 아니라 후천적 요인인 환경에 대한 적응, 식습관, 운동 등과 같은 요소로도 신장, 체형, 질병, 노화 등 건강 전 반에 걸쳐 영향을 주게 된다[5-7].

이를 한마디로 “염기서열을 넘어서(beyond the DNA sequence)”라고 표현할 수 있으며, 이에 대한 연구를 후성유전학(epigenetics)라고 한다[7,8]. 후성유전 조절 은 세포 분화를 포함한 다양한 조절반응을 이해할 수 있는 분자 메커니즘으로 부각되고 있으며, 세포의 기 능 분석 및 질환의 병인 규명 연구에 필수적이다. 생명 과학적 측면에서 후성유전 연구는 크게 세 분야에서 접근할 수 있는데, DNA 메틸화, 히스톤의 변화, non-coding RNA가 그것이다(그림 1).

• DNA 메틸화(DNA methylation)

DNA 메틸화는 가장 잘 알려진 화학적 변형 중 하나이다. 포유동물에서 DNA의 메틸화는 CpG 염기서열에 있는 “C(시토신)”에 메틸화를 시키는 과정을 일컫는다. 이 변화는 유전자의 발현 억제, 염색체의 안정성, 유전자 각인 등의 유전체 기능 에 매우 중요하게 작용한다. 인간 게놈에 CpG가

나올 확률은 6.25%(1/16)인데 인간 게놈의 98% 에는 대략 1~2%에 지나지 않는다. 이를 CpG 억 제(CpG suppression)라고 한다[9]. 이에 반해 특정 지역에 CpG가 확률 이상으로 발견되는데 이를 CpG island라고 하는데 이곳에 있는 CpG의 약 80%는 cytosine 염기의 5번 위치에 메틸화되어 있기에 “제 5의 염기”라고도 불린다[10]. 대부분 사람 유전자의 프로모터에는 CpG island가 존재 하며 이 CpG에 있는 시토신 염기가 메틸화되면 전사인자의 접근이 차단되어 유전자가 억제된다 [11].

• 히스톤의 변화

뉴클레오솜(nucleosome)은 146bp 길이의 DNA가 8개로 구성된 복합체인 히스톤 단백질에 감겨있 는 DNA 저장의 가장 작은 단위이다[12]. 이 히 스톤 N-말단부위에는 아미노산이 아세틸화, 메틸 화, 유비퀴틴화, 인산화 등과 같은 화학적 변형이 많이 일어난다. 이는 염색체의 전반적인 구조의 변화를 야기함으로써 DNA-히스톤의 결합에 영 향을 미치기 때문에 유전자의 발현을 조절하게 된다[13].

• non-coding RNA

단백질로 번역이 되지 않는 non-coding RNA는 tRNA, rRNA를 비롯하여 small nucleolar RNA(sno-RNA), microRNA(miRNA), short-interfering RNA (siRNA) 등이 있는데 이 중 유전자 조절 기능으로 잘 알려진 것은 miRNA이다. 타깃이 되는 유전자 에 결합하여 전사를 억제하는 방식으로 유전자 조 절에 관여하고 있다[14]. 약 50%의 단백질을 코딩 하는 유전자가 miRNA의 직접적인 타깃이 되고 그 이외에도 간접적 영향을 미치는 유전자 타깃 을 고려하면 세포내에서 일어나는 거의 모든 일 과 관계가 있을 것으로 추산한다.

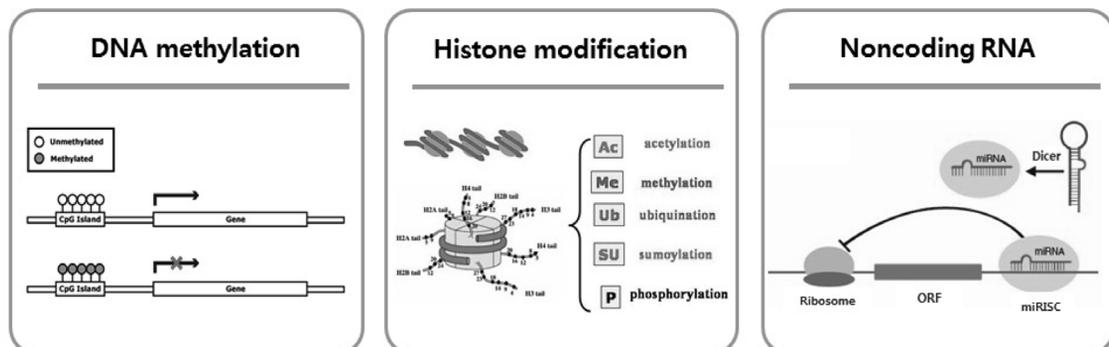


그림 1 후성유전학의 주요 연구 분야[15-17]

miRNA는 짧은 non-coding RNA로 유전자 발현을 조절할 수 있는 메커니즘 중 하나이다. 이 miRNA들은 경우에 따라 인간 전체유전자의 30%~92%를 조절하기 때문에, 세포증식(proliferation)과 발달(development), 사멸(apoptosis), 종양형성(tumorigenesis)과 형태발생(morphogenesis), 조혈과정(hematopoiesis)과 면역세포발달 및 분화 등을 포함한 거의 모든 생명현상에 관여함으로써 세포 항상성을 유지한다[18-23]. 따라서 이들 miRNA의 작용에 문제가 생길 경우 부적절한 세포분화와 증식, 세포사멸을 유도해 암이나 자가면역질환 같은 질병을 유발 할 수 있다.

miRNA의 기작에 대한 이해는 새로운 질병 진단 마커(diagnostic marker)뿐만 아니라 miRNA을 조절함으로써 질병치료에도 사용될 수 있을 것이다. 이를 위해서는 miRNA의 타겟이 되는 유전자와 그 기작에 대한 이해가 필수적이다. miRNA의 기능은 대개 miRNA가 조절하는 유전자를 통하여 알 수 있지만 조절 체계가 매우 복잡하여 이해하기가 쉽지 않다. 이에 본 리뷰에서는 miRNA에 대해 알아보고 miRNA의 타겟이 되는 유전자를 찾는 알고리즘에 대해 정리해보고자 한다.

2. 후성유전학적 조절인자, miRNA

miRNA는 21-25개의 뉴클레오타이드(nucleotide, nt)로 이루어진 단일 염기가닥의 small non-coding RNA로서, 식물과 동물 모두에서 발견되는 유전자 조절하는 중요한 조절인자다[24,25]. miRNA가 발견되기 전까지의 RNA의 역할은 DNA에 담겨있는 유전정보를 단백질로 변환시키는 과정에서 단순히 정보를 전달하는 역할로만 인식되어 왔다. DNA 염기서열은 mRNA로 전사되고 세포질의 리보솜의 tRNA를 거쳐 단백질로

합성된다. 그런데 miRNA는 이 과정에서 mRNA와 상보적인 결합을 통해 세포내 유전자 발현을 조절하는 중추적 역할을 하는 것이다(그림 2)[26].

miRNA의 생성은 두 단계를 통해 이루어지는데, 핵 안에서 DNA가 RNA polymerase II에 의해 전사되어 pri-miRNA라 불리는 비교적 큰 RNA 전구체(large RNA precursor)가 만들어진다. pri-miRNA는 mRNA와 매우 유사한 구조로 전사되는데[27], pri-miRNA는 핵 속에 존재하는 RNAase III 단백질인 Drosha에 의해 절단된 후, 약 70개의 염기로 구성된 불완전한 stem-loop 구조인 pre-miRNA를 만든다[28]. pre-miRNA는 RNA-GTP 의존성 transporter에 의해서 세포질로 나온 뒤 세포질에 존재하는 또 다른 RNAase III 단백질인 Dicer의 작용에 의해 헤어핀(hairpin)구조가 절단되어 약 22개의 염기로 구성된 최종 miRNA(mature miRNA)가 생성된다[29]. 이 후 miRNA는 Argonaute(Ago)를 포함하여 여러 가지 단백질과 함께 RNA-induced silencing complex(RISC)라고 불리는 복합체를 형성하고[30], 이 복합

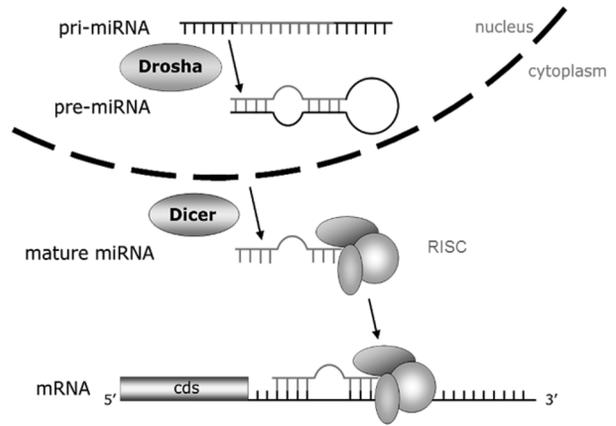


그림 3 miRNA 만들어지는 과정[32]

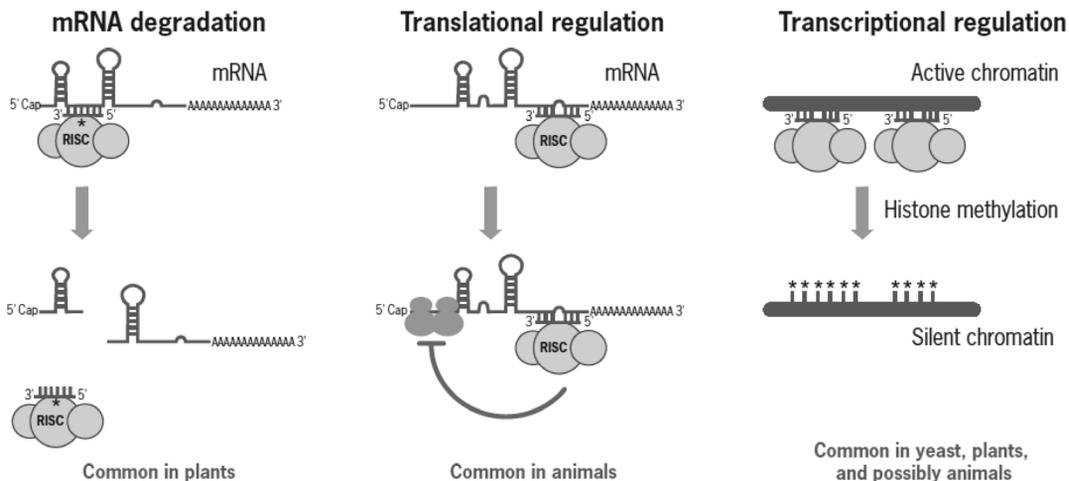


그림 2 miRNA의 유전자 조절(<http://www.ambion.com>)

체가 표적 유전자와 결합함으로써 표적 mRNA를 분해시키거나 단백질로 번역되는 과정을 억제하게 된다[31].

3. 생물정보학 기반 miRNA 타겟 예측

miRNA의 기능은 대개 miRNA가 조절하는 유전자를 통하여 알 수 있다. 이들 miRNA는 2013년 1월 기준 human miRNA 2,578개(mature miRNA)가 데이터베이스화 되어 있다(miRBase database, version 20.0). 수백~수천 개의 타겟 mRNA들의 3' untranslated region(UTR)에 상보적으로 결합해 타겟 유전자의 발현을 억제한다. 그러나 포유동물의 경우에는 타겟 유전자와 불일치염기쌍(mismatched base-pair)과 G-U 염기쌍을 허용하는 불완전한 상보적 결합을 하고 서열의 길이가 짧기 때문에 기존의 서열 비교 방법(BLAST)으로 타겟 유전자를 찾는 것은 쉬운 일이 아니다. 최근에는 생물학적 타겟 유전자의 3'UTR과 miRNA의 seed를 alignment하여 예측하는데 그치지 않고 구조적인 특징을 덧붙이거나, 진화적으로 보존된 결합자리 등을 추가함으로써 타겟 유전자 예측의 정확도를 높이고 있다[33].

이러한 예측 결과를 제공하는 프로그램으로는 DIANA-microT[34], miRanda [35], MirTarget2[36], miTarget [37], PicTar[38], PITA[39], and TargetScan[40] 등이 있다(표 1). 그러나 이러한 프로그램들은 miRNA마다 수십에서 수백 개의 타겟 유전자를 조절한다고 예측하기 때문에 실험을 통하거나 다른 여과 인자를 추가하여 false positive를 제거하는 과정이 필수적이다[41,42].

생물정보학적 방법으로 miRNA의 타겟을 찾을 때 주로 사용되는 알고리즘들은 서열과 MRE(miRNA recognition elements) 위치를 바탕으로 진행이 되는데 특

히 아래의 4가지 요인에 가중치를 두고 있다[44].

- miRNA seed region
- Evolutionary conservation(진화적 보존성) of the MRE
- Free energy(자유에너지, ΔG) of the miRNA-mRNA hetero duplex
- miRNA sequence features outside the target site

3.1 Seed 중심 방법

TargetScan, PicTar의 경우 seed region 중심이다. TargetScan의 경우 seed region이 최소 7 염기 이상 완전 결합이 돼야 하는 반면에, PicTar는 불완전 결합을 용인하면서 자유에너지 값으로 제거시킨다. miRanda는 miRNA가 타겟 mRNA와 얼마나 높은 비율로 상호 결합하느냐에 착안을 하는데 특히 G·U wobble과 mismatch와 같은 결합보다는 seed 결합에 높은 가중치를 부여한다. 하지만 seed와 완전결합을 요구하지는 않는다[45].

각 알고리즘에 대해 좀 더 자세히 살펴보면, TargetScan은 여러 종에서 보존된 7-8mer의 miRNA sites를 모아 miRNA의 family로 모티프(motif)를 만들어 서로 상보적으로 결합할 수 있는 후보 서열들 mRNA의 3'-UTR 서열로부터 결합자리와 match되는 유전자 타겟을 예측하였다. 그러나 종간의 보존도가 있는 miRNA와 mRNA만을 사용했기 때문에 아직 다른 종에서 발견되지 않은 miRNA는 데이터로 사용되기 힘들다. 반면 miRanda의 경우에는 이들의 상보성뿐만 아니라 결합에 의한 자유에너지를 계산한 후 가장 안정한 결합 구조를 가지는 지의 여부를 계산하여 표적 유전자를 예측한다. PicTar에 적용되는 알고리즘은 표적 유전자

표 1 miRNA 타겟 예측 프로그램 [43]

Method	Type of Method	Organism	Resource
miRanda	Complementary	Flies, vertebrates	http://www.microrna.org
TargetScan	Seed complementary	Vertebrates	http://www.targetscan.org
DIANA-microT	Thermodynamics		http://diana.cslab.ece.ntua.gr/
PicTar	Thermodynamics	Vertebrates, flies, nematodes	http://pictar.mdc-berlin.de/
PITA	Thermodynamics		http://genie.weizmann.ac.il/pubs/mir07/index.html
RNAHybrid	Thermodynamics & Statistical model	Flies	http://bibiserv.techfak.uni-bielefeld.de/rnahybrid
miRGen++	Baynesian Inference		http://www.psi.toronto.edu/genmir
MiTarget	Support Vector Machine		http://cbt.snu.ac.kr/~miTarget
MiRtarget2	Support Vector Machine	Vertebrates	http://mirdb.org
TarBase	Experimentally Validated Targets		http://diana.cslab.ece.ntua.gr/tarbase/

의 3'UTR과 miRNA의 서열 간에 다중 서열비교(multiple sequence alignment)를 통해 서열의 유사성을 통계 기법을 이용한 확률 값으로 계산하여 표적 유전자를 예측하는 알고리즘을 사용한다[31].

3.2 타깃 자리 접근성(target site accessibility)

이 알고리즘은 miRNA와 타깃간의 duplex 형성으로 얻을 수 있는 자유에너지 에너지만을 고려하는 것이 아니라 miRNA와 성공적으로 결합할 수 있는 타깃의 2차 구조를 고려하여 실제 miRNA가 타깃자리에 접근가능한가($\Delta\Delta G$)를 계산한다. PITA는 사용자가 허용하는 mismatch가 포함된 seed 부위를 찾고 각 match된 것마다의 $\Delta\Delta G$ 를 계산한다. 즉, 3'UTR에서 각 miRNA의 initial seed를 가지고 model에 적용하여 예상 site를 찾고, 같은 miRNA의 site들은 합하여 miRNA와 UTR의 총 상호작용 점수(total interaction score)를 얻는다. 비록 3'UTR에 적용 가능하도록 디자인된 알고리즘이지만 3'UTR 이외의 부분에 있는 MRE도 예측가능하다[42].

3.3 miRNA sequence pattern

다른 알고리즘과는 다르게 rna22 알고리즘은 공유하는 miRNA 서열 패턴을 기반으로 이미 알려져 있는 mature miRNA 서열을 training dataset으로 하여 통계적으로 의미 있는 패턴을 찾게 된다. 가능성 있는 MRE는 상호보완으로 결합하는 부위로 찾고 자유에너지 값을 고려한다. 특히, ma22는 각 종마다 보존되는 부위

에 대해 고려하지 않고, 3'UTR에 한정짓지도 않으며 seed의 mismatch를 허용한다. rna22의 개발자들은 낮은 false-positive를 주장하지만, small scale 데이터 기반으로 예측되었으므로 high-throughput 데이터 기반으로 만들어진 seed 중심 방법과는 차이가 있을 것으로 예측된다[44,46].

4. miRNA 타깃 예측 알고리즘의 비교

위에서 소개한 예측 결과들은 miRNA당 예측한 타깃 유전자의 개수가 수십에서 수백 개에 이르며, false-positive 데이터와 같은 오류도 많기 때문에 결과에 대한 실험적 검증이 필요하다. 예측 결과들 중 동일한 miRNA에 대하여 타깃을 예측한 것들이므로 동일하게 예측한 결과들 또한 존재한다. 여러 예측 프로그램이 동일하게 예측한 결과는 좀 더 정확도가 높을 것이라 예상하고 각 프로그램별로 비교해보았다.

본 집필진은 인간 데이터베이스를 기준으로, PicTar의 경우 4개의 척추동물을 이용한 PicTar_4way 파일을 이용하였으며, Miranda와 TargetScan은 conserved한 부위를 중심으로 각기 Miranda_conserved, TargetScan_Conservedsite 파일을 이용하였고, PITA는 flanking한 염기서열을 upstream에서 3 nucleotide(nt), downstream으로 5 nt까지 허용하고 seed에서 보존적인 7- 또는 8-mer를 가진 PITA_targets_hg18_3_5_TOP 파일을 사용하였다(표 3).

표 2 miRNA target 예측 알고리즘 비교[44]

MRE feature	Example	Target prediction algorithm*				
		TargetScan	PicTar	miRanda	PITA	ma22
Perfect seed	<p>HMGA2 5' CCGACAUUCAAUUUCUACCCUCA 3'</p> <p>let-7a 3' UUGAUUUGUUGGAUGAUGGAGU 5'</p>	-				
G:U wobble seed	<p>NF2 5' UACAAGAGAUUCUCUGCCUCA 3'</p> <p>let-7a 3' UUGAUUUGUUGGAUGAUGGAGU 5'</p>	-	-			
Imperfect seed/seedless	<p>E2F2 5' GUGGUGUCU-CUGGGCUGAACCA 3'</p> <p>miR-24 3' GACA-AGGACGACUUGACUCGGU 5'</p>	-	-			
Outside the 3' UTR	<p>DNMT3B 5' UGGCAAAGAAGAUUUUUGUGGUGCAGAG 3'</p> <p>miR-148 3' -UGUUU.....CAAGACAUCACGUGACU- 5'</p>	-	-	-	+	+

표 3 miRNA target database status

Predicted miRNA Target Database	miRNA	Target	miRNA-Target	Database version
PicTar	178	6,754	55,893	PicTar_4way
Miranda	241	19,271	708,229	Miranda_conserved
PITA	677	10,143	208,937	PITA_targets_hg18_3_5_TOP
TargetScan	1,537	15,031	520,354	TargetScan_Conservedsite
Total	1,603	21,394	1,221,262	

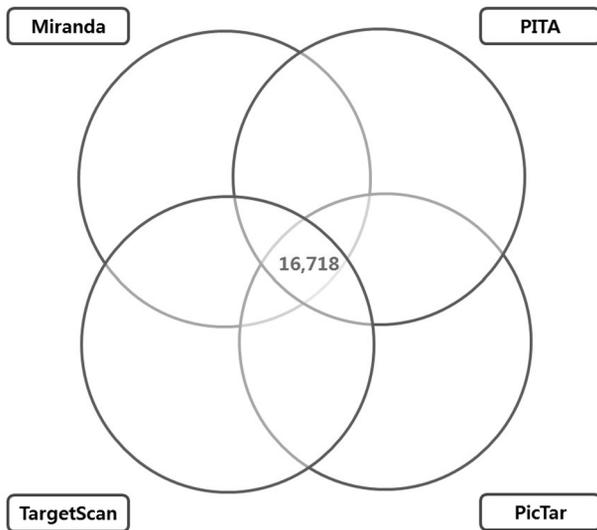


그림 4 miRNA-mRNA pair 중복 비교

각 데이터베이스에서 제공하는 miRNA와 타겟 mRNA의 수 및 miRNA-타겟 쌍의 수를 조사한 결과, 4개 데이터베이스 모두에서 중복되어 나타나는 miRNA-target pair는 16,718개(122 miRNA, 3,232 target)로 나타났다. 가장 많은 수의 pair 정보를 제공하는 Miranda 기준으로 약 2.3%를, PicTar 기준으로 약 30%에 해당하는 miRNA-target pair의 정보만이 정확도가 높다고 말할 수 있는 것이다(그림 4). 게다가 4개 데이터베이스에서 모두 중복되어 나타나는 pair의 수는 4개 데이터베이스를 모두 합친 것 기준으로 1.4%의 pair 정보만을 신뢰할 수 있다는 말이 되므로 그 수가 매우 적다고 말할 수 있다.

이와 같은 결과는 각 사이트별로 miRNA의 ID가 정리되어 있다고는 하나 mature miRNA와 헤어핀(hair-pin) miRNA 등의 ID를 혼재해서 쓰는 등의 문제가 존재하고 있기 때문으로 보인다. 위의 예측 프로그램의 결과에 대한 보다 정확한 평가는 실험적 검증이 이루어져야 하며 그 후 추가적인 인자를 보완하여 예측하는 진화 알고리즘이 향후 개발되어야 할 것이다.

5. 결론

분자생물학, 세포생물학, 제약 및 의료 분야에서 기존에 밝혀진 바이오 정보를 바탕으로 구축되어진 알고리즘을 통해 예측된 결과의 정확도가 점점 높아짐에 따라 생물정보학에 대한 관심이 고조되고 있는 것이 현실이며 고무적이다. 또한, high-throughput 기술의 발달로 쏟아져 나오는 데이터를 정보과학 기술과 접목시켜 생명 현상의 비밀을 밝히는데 필수적인 기술이자 학문으로 대두되고 있다. 본고에서는 최근에

빠른 속도로 연구가 진행되고 있는 후성유전학 기작들 중 miRNA와 그 타겟을 선정하는 알고리즘에 대해 알아보았다. miRNA 타겟 예측 알고리즘은 seed 서열, 2차 구조를 이용한 접근성 및 miRNA 패턴을 이용하였다. miRNA 타겟 예측 데이터베이스간 중복되는 miRNA-target 수가 매우 적은데, 이는 각 데이터베이스별 중심으로 다루는 원리에 차이가 있기 때문이다. 또한, 알고리즘 기반으로 예측된 타겟에는 많은 false-positive 데이터를 포함하고 있어 실제 타겟과 100% 결과가 옳다고 할 수는 없지만, 실험 전 예측된 타겟을 이용하여 범위를 줄이고 실험에 들어가는 노력, 돈, 시간을 단축한다는 데 크게 기여를 하고 있다. 대량의 유전자 발현정보와 NGS의 결과를 현재 알고리즘과 결합하려는 노력이 많이 시도되고 있으며 좀 더 정확도가 높은 결과를 예측하고자 노력하고 있다[47].

예측결과와 실제 결과간의 간극을 줄이기 위해서는 정보기술과 생명과학 간의 긴밀한 연구 개발의 공조 체제를 구축하여야 할 것이다. 특히, 정보기술 기반 생명과학자나 생명과학 기반 정보과학자들이 중심이 되어 융합학문으로의 새로운 도약이 필요한 시점이다. 이들이 서로 공유하고 협력하는 차세대 연구 개발자들의 출현이 기다려지는 이유는 그 동안 생물정보학이 빠르게 달려왔지만 앞으로 달려갈 길이 더 많이 남아 있기 때문이다.

참고문헌

- [1] 원세연, “생물정보학이란 무엇인가”, 과학기술정책, 9월호, pp. 71-78, 2000.
- [2] Jones, P. B. C., “The commercialization of bioinformatics”, Electronic Journal of Biotechnology, Vol. 3, No. 2, 2000.
- [3] Kitano, H., “Computational systems biology”, Nature, Vol. 420, No. 6912, pp. 206-210, 2002.
- [4] Stemple, D. L., “TILLING--a high-throughput harvest for functional genomics”, Nat Rev Genet, Vol. 5, No. 2, pp. 145-150, 2004.
- [5] Venter, J. C., Adams, M. D., Myers, E. W., Li, P. W., Mural, R. J., Sutton, G. G., Smith, H. O., Yandell, M., Evans, C. A., Holt, R. A. et al, “The sequence of the human genome”, Science, Vol. 291, No. 5507, pp. 1304-1361, 2001.
- [6] Lander, E. S., Linton, L. M., Birren, B., Nusbaum, C., Zody, M. C., Baldwin, J., Devon, K., Dewar, K., Doyle, M., FitzHugh, W. et al, “Initial sequencing and analysis of the human genome”, Nature, Vol. 409, No. 6822, pp.

860-921, 2001.

- [7] Jaenisch, R. and Bird, A., "Epigenetic regulation of gene expression: how the genome integrates intrinsic and environmental signals", *Nat. Genet.*, Vol. 33, pp. 245-254, 2003.
- [8] Wolffe, A. P. and Matzke, M. A., "Epigenetics: regulation through repression", *Science*, Vol. 286, No. 5439, pp. 481-486, 1999.
- [9] Gardiner-Garden, M. and Frommer, M., "CpG islands in vertebrate genomes", *J. Mol. Biol.*, Vol. 196, No. 2, pp. 261-282, 1987.
- [10] Watt, F. and Molloy, P. L., "Cytosine methylation prevents binding to DNA of a HeLa cell transcription factor required for optimal expression of the adenovirus major late promoter", *Genes Dev.*, Vol. 2, No. 9, pp. 1136-1143, 1988.
- [11] Boyes, J. and Bird, A., "DNA methylation inhibits transcription indirectly via a methyl-CpG binding protein", *Cell*, Vol. 64, No. 6, pp. 1123-1134, 1991.
- [12] Richmond, T. J. and Davey, C. A., "The structure of DNA in the nucleosome core", *Nature*, Vol. 423, No. 6936, pp. 145-150, 2003.
- [13] Spencer, V. A. and Davie, J. R., "Role of covalent modifications of histones in regulating gene expression", *Gene*, Vol. 240, No. 1, pp. 1-12, 1999.
- [14] Valencia-Sanchez, M. A., Liu, J., Hannon, G. J., and Parker, R., "Control of translation and mRNA degradation by miRNAs and siRNAs", *Genes Dev.*, Vol. 20, No. 5, pp. 515-524, 2006.
- [15] http://missinglink.ucsf.edu/lm/genes_and_genomes/methylation.htm.
- [16] <http://chemistry.gsu.edu/faculty/Zheng/pictures/nucleosome.jpg>.
- [17] 노태영, "후성유전체와 전사조절", *Bioin 스페셜*, 13호, 2009.
- [18] Ambros V., "MicroRNA pathways in flies and worms: growth, death, fat, stress, and timing", *Cell*, Vol. 113, No. 6, pp. 673-676, 2003.
- [19] Chen, C. Z., Li, L., Lodish, H. F., and Bartel, D. P., "MicroRNAs modulate hematopoietic lineage differentiation", *Science* Vol. 303, No. 5654, pp. 83-86, 2004.
- [20] Xu, P., Vernoooy, S. Y., Guo, M., and Hay, B. A. "The *Drosophila* microRNA Mir-14 suppresses cell death and is required for normal fat metabolism", *Curr Biol*, Vol. 13, No. 9, pp. 790-795, 2003.
- [21] Pfeffer, S., Zavolan, M., Grasser, F. A., Chien, M., Russo, J. J., Ju, J., John, B., Enright, A. J., Marks, D., Sander, C., and Tuschl, T., "Identification of virus-encoded microRNAs", *Science*, Vol. 304, No. 5671, pp. 734-736, 2004.
- [22] Calin, G. A., Sevignani, C., Dumitru, C. D., Hyslop, T., Noch, E., Yendamuri, S., Shimizu, M., Rattan, S., Bullrich, F., Negrini, M., and Croce, C. M., "Human microRNA genes are frequently located at fragile sites and genomic regions involved in cancers", *Proc Natl Acad Sci USA*, Vol. 101, No. 9, pp. 2999-3004, 2004.
- [23] Calin, G. A., Dumitru, C. D., Shimizu, M., Bichi, R., Zupo, S., Noch E., Aldler, H., Rattan, S., Keating, M., Rai, K., Rassenti, L., Kipps, T., Negrini, M., Bullrich, F., and Croce, C. M., "Frequent deletions and down-regulation of micro-RNA genes miR15 and miR16 at 13q14 in chronic lymphocytic leukemia", *Proc Natl Acad Sci USA*, Vol. 99, No. 24, pp. 15524-15529, 2002.
- [24] Reinhart, B. J., Weinstein, E. G., Rhoades, M. W., Bartel, B., and Bartel, D. P., "MicroRNAs in plants", *Genes Dev.*, Vol. 16, No. 13, pp. 1616-1626, 2002.
- [25] Stefani, G. and Slack, F. J., "Small non-coding RNAs in animal development", *Nat Rev Mol Cell Biol.*, Vol. 9, No. 3, pp. 219-230, 2008.
- [26] He, L. and Hannon, G. J., "MicroRNAs: small RNAs with a big role in gene regulation", *Nature Reviews Genetics*, Vol. 5, No. 7, pp.522-531, 2004.
- [27] Cai, X., Hagedorn, C. H., and Cullen, B. R., "Human microRNAs are processed from capped, polyadenylated transcripts that can also function as mRNAs", *RNA*, Vol. 10, No. 12, pp. 1957-1966, 2004.
- [28] Lee, Y., Ahn, C., Han, J., Choi, H., Kim, J., et al. "The nuclear RNase III Droscha initiates microRNA processing", *Nature*, Vol. 425, No. 6956, pp. 415-419, 2003.
- [29] Saito, K., Ishizuka, A., Siomi, H., and Siomi, M. C., "Processing of pre-microRNAs by the Dicer-1-Loquacious complex in *Drosophila* cells", *PLoS Biol.*, Vol. 3, No. 7, e235, 2005.
- [30] Gregory, R. I., Chendrimada, T. P., Cooch, N., and Shiekhattar, R., "Human RISC couples microRNA biogenesis and posttranscriptional gene silencing", *Cell*, Vol. 123, No. 4, pp. 631-640, 2005.
- [31] 이광희, "마이크로RNA의 생물정보학적 분석 및 연구방법", *질병관리본부 주간건강과 질병*, 제4권, 제16호, pp. 283-287, 2011.
- [32] Brzuzan, P., Wozny, M., Wolinska, L. and Luczynski,

- M. K., An Integrated view of the molecular recognition and toxinology-From analytical procedures to biomedical applications, INTECH, pp. 223-237, 2013.
- [33] Alexiou, P., et al., "Lost in translation: an assessment and perspective for computational microRNA target identification", *Bioinformatics*, Vol. 25, No. 23, pp. 3049-3055, 2009.
- [34] Kiriakidou, M., Nelson, P. T., Kouranov, A., Fitziev, P., Bouyioukos, C., Mourelatos, Z., and Hatzigeorgiou, A., "A combined computational-experimental approach predicts human microRNA targets", *Genes Dev.*, Vol. 18, No. 10, pp. 1165-1178, 2004.
- [35] John, B., Enright, A. J., Aravin, A., Tuschl, T., Sander, C., and Marks, D. S., "Human MicroRNA targets", *PLoS Biol.*, Vol. 2, No. 11, pp. e363, 2004.
- [36] Wang, X. and El Naqa, I. M., "Prediction of both conserved and nonconserved microRNA targets in animals", *Bioinformatics*, Vol. 24, No. 3, pp. 325-332, 2008.
- [37] Kim, S. K., Nam, J. W., Rhee, J. K., Lee, W. J., and Zhang, B. T., "miTarget: microRNA target gene prediction using a support vector machine", *BMC Bioinformatics*, Vol. 7, pp. 411, 2006.
- [38] Krek, A., Grün, D., Poy, M. N., Wolf, R., Rosenberg, L., Epstein, E. J., MacMenamin, P., da Piedade, I., Gunsalus, K. C., Stoffel, M., and Rajewsky, N., "Combinatorial microRNA target predictions", *Nat Genet.* Vol. 37, No. 5, pp. 495-500, 2005.
- [39] Kertesz, M., et al., "The role of site accessibility in microRNA target recognition", *Nat Genet.* Vol. 39, No. 10, pp. 1278-1284, 2007.
- [40] Lewis, B. P., Shih, I. H., Jones-Rhoades, M. W., Bartel, D. P., and Burge, C. B., "Prediction of mammalian microRNA targets", *Cell*, Vol. 115, No. 7, pp. 787-798, 2003.
- [41] Croce, C.M., "Causes and consequences of microRNA dysregulation in cancer", *Nat Rev Genet.* Vol. 10, No. 10, pp. 704-714, 2009.
- [42] 서채화, "Prediction of microRNA relevance to cancer based on gene set analysis", 석사학위논문, 이화여자대학교 대학원, 서울, 2009.
- [43] Mazière, P., Enright, A. J., "Prediction of microRNA targets", *Drug Discov Today*, Vol. 12, No. 11-12, pp. 452-458, 2007.
- [44] Thomas, M., Lieberman, J., and Lal, A., "desperately seeking microrna targets", *Nat Struct Mol Biol.* Vol. 17, No. 10, pp. 1169-1174, 2010.
- [45] Betel, D., Wilson, M., Gabow, A., Marks, D. S., and Sander, C., "The microRNA.org resource: targets and expression", *Nucleic Acids Res.*, Vol. 36, pp. 149-153
- [46] Miranda, K. C., Huynh, T., Tay, Y., Ang, Y. S., Tam, W. L., Thomson, A. M., Lim, B., and Rigoutsos, I., "A pattern-based method for the identification of MicroRNA binding sites and their corresponding heteroduplexes", *Cell*, Vol. 126, No. 6, pp. 1203-1217, 2006.
- [47] Nelson, P. T., Baldwin, D. A., Scearce, L. M., Oberholzer, J. C., Tobias, J. W., and Mourelatos, Z., "Microarray-based, high-throughput gene expression profiling of microRNAs", *Nat Methods*, Vol. 1, No. 2, pp.155-161, 2004.

약 력



류 제운

2006 충북대학교 생화학과(이학사)
 2008 충북대학교 생화학과(이학석사)
 2010~현재 충북대학교 생화학과 박사과정
 관심분야: 생물정보학, 단백질 네트워크, 후성 유전학
 E-mail : jwryu84@cbnu.ac.kr



유 재수

1989 전북대학교 컴퓨터공학과(공학사)
 1991 KAIST 전산학과(공학석사)
 1995 KAIST 전산학과(공학박사)
 1996~현재 충북대학교 정보통신공학부 및 컴퓨터정보통신연구소 교수
 관심분야: 생물정보학, 신호전이, 시스템 생물학,

후성유전학
 E-mail : yjs@cbnu.ac.kr



김 학용

1985 충북대학교 농화학과(농학사)
 1987 충북대학교 화학과(이학석사)
 1994 미국 코네티컷대학교 분자·세포생물학과(이학박사)
 1998~현재 충북대학교 생화학과 교수
 관심분야: 시스템생물학, 신호전이, 단백질 네트

워크, 생체동역학
 E-mail : hykim@cbnu.ac.kr