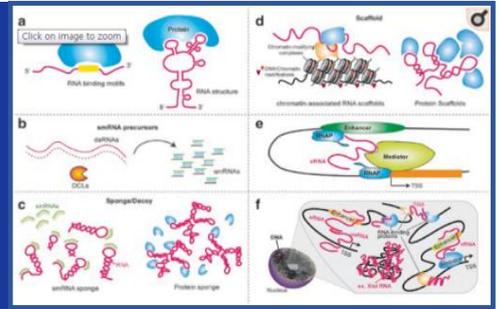


Long non-coding RNAs (LncRNAs) 분석

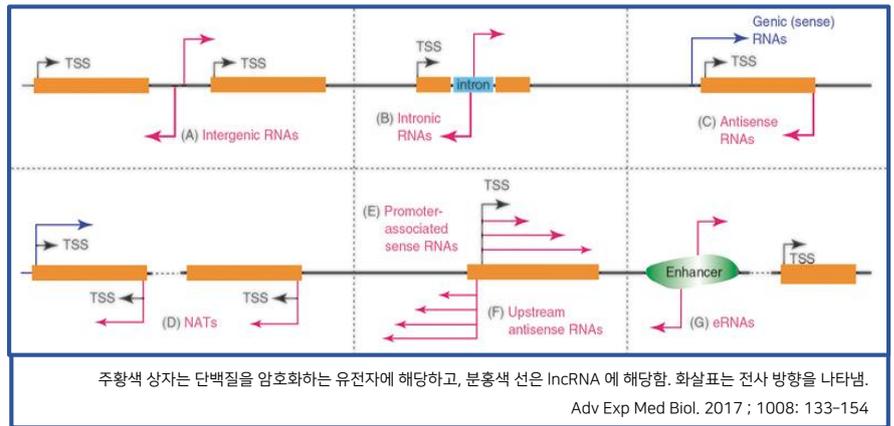
추가 시퀀싱 없이, 기존 RNAseq 데이터를 이용하여 gene expression 분석, LncRNAs 분석을 동시에 진행 가능!



Non-coding RNA (ncRNA) 는 DNA 에서 전사되지만, 단백질은 형성되지 않는 기능을 갖는 RNA 분자로, 비록 단백질을 생성하지는 못하지만 최근 들어 많은 연구가 진행되고 있는 전사 후 과정 (Post-transcriptional regulation)과 후성유전학 (Epigenetics) 적 메커니즘들을 작동시키는 주요한 요인들을 제어하는 역할을 하는 것으로 밝혀지며 큰 관심을 받고 있다. ncRNA 들은 물리적 길이 (200 nucleotide)에 따라 small non-coding RNA (sncRNAs) 와 long noncoding RNAs (lncRNAs) 로 나뉘어 진다.

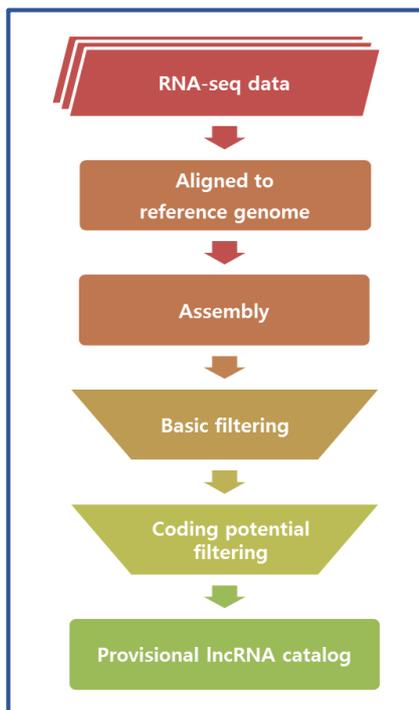
LncRNAs 의 분류

lncRNAs 는 단백질을 암호화한 유전자의 유전체 내 구조 관계를 기초로 분류된다. 오른쪽의 그림은 7가지 종류의 lncRNAs: intergenic or long intergenic noncoding RNAs (lincRNAs) (A), intronic RNAs (B), antisense RNAs (C), natural anti-sense transcripts (NATs) (D), promoter-proximal sense (E) and upstream antisense RNAs (F), eRNAs (G)를 묘사한 그림이다.



NGS 를 이용한 LncRNAs 분석

분석파이프라인



lncRNAs 분석은 각 샘플 별 RNAseq 데이터를 이용하여 genome-assisted assembly 또는 *de novo* assembly 를 통해 transcripts model 을 확보하고 하나로 통합한 뒤에 단계별 필터과정을 거친다.

필터는 크게 두 단계로 나뉘는데, 먼저 basic filtering 과정으로 어셈블러나 샘플 별 반복성을 가지며 어느정도 read coverage를 갖는 신뢰할 수 있는 transcripts를 선별한다. 또한 상단의 그림처럼 유전체 내에서 lncRNA의 구조를 가지거나 알려진 단백질 서열과 유사성이 없는 transcripts를 후보로 선별한다. 두 번째 필터는 coding potential filtering 과정으로, CPC(<http://cpc.cbi.pku.edu.cn/>)와 같은 coding 가능성을 계산해주는 tool을 이용하여 단백질 생성 가능성이 없는 transcripts만을 최종 lncRNA 후보로 선별한다. 이 결과를 이용하여 샘플 별 lncRNA의 발현 분석이나 응용 분석을 진행할 수 있다.

분석포인트!!!

lncRNAs 는 염기서열의 길이가 200 nucleotide 이상이면서 mRNA 처럼 말단에 Adenine 염기만 중복된 긴 꼬리 (poly(A) tail)를 갖지만, 단백질 합성에 실제로 이용될 가장 중요한 염기서열부분인 open reading frame (ORF) 가 없는 비암호화 RNA 에 한 부류 이므로, 새로운 lncRNA 를 찾은 이후에는 단백질 생성가능성에 대한 검토가 반드시 필요하다. 단백질 생성가능성이 높은 도메인을 포함하는지 여부 또는 단백질 생성을 위해 잘 보존된 아미노산 함유도를 분석하여 단백질 생성가능성을 검토한다.

Contact Us

본사 042.710.4035 | www.seeders.co.kr | master@seeders.co.kr | 대전시 중구 중앙로 118, 5층 501호

문의 사항은 위 연락처로 연락 주시기 바랍니다. For further details of our technology, please contact SEEDERS.

