

기타 분석 도구

- 기타 분석 도구 메뉴는 시각화 기능 외 생물학적 서열 데이터(FASTA) 또는 Variant Call Format(VCF) 파일을 분석하기 위한 메뉴입니다.
- 해당 메뉴는 생물학적 서열 데이터를 분석하는 **BLAST**, **Interpro**, VCF 파일을 MAF 파일로 변환 해주는 **VCFtoMAF**, 여러 개의 MAF 파일을 하나의 파일로 병합시켜주는 **MAF Merger**, VCF 파일의 참조 유전체 버전을 변환(hg19 ↔ hg38) 시켜주는 **RefVer Converter**로 구성됩니다.

사이트 소개 데이터 분석 예시 **사용자 데이터 분석** 고객의 소리

사용자 데이터 분석

홈 > 사용자 데이터 분석 > 기타 분석 도구

기타 분석 도구

BLAST
Basic Local Alignment Search Tool

Blast
생물학적 서열 데이터를 입력하여 database에서 유사 서열 검색

mskcc/vcf2maf
Convert a VCF into a MAF, where each variant is associated to only one of all possible gene isoforms.

VCF To MAF
VCF 파일을 MAF 형식으로 변환

InterPro
Classification of protein families

Interpro
사용자의 단백질 서열을 입력하여 interpro 데이터베이스로부터 domain 및 family 분석

MAF Merger
여러 개의 MAF 파일을 입력하여 1 개 파일로 병합하는 분석 도구

RefVer Converter (Liftover)
VCF 파일의 참조 유전체 버전을 변환 (hg19 ↔ hg38)

Dataframe Converter
매트릭스 형태의 파일을 행렬 형태로 변환

< 사용자 데이터 분석 - 기타 분석 도구 화면 >

BLAST

- BLAST는 Basic Local Alignment Search Tool의 약자로 생물학적 서열 데이터(FASTA 파일)를 위한 분석 도구입니다.
- ① 서열 입력 상자에 직접 FASTA 서열 정보를 입력하거나 또는 FASTA 파일을 Drag & Drop 합니다.
- ② 본인의 서열 데이터가 염기서열 기반이라면 Nucleotide databases를 선택하고, 아미노산 기반이라면 Protein databases를 선택합니다.
- ③ Advanced parameters 상자에서 parameter를 설정 ([?] 박스를 누르면 parameter 종류 확인 가능)
예시) -evalue 1.0e-5 -num_alignments 100
(입력 정보가 없으면 default 값으로 설정. 특별한 경우가 아니면 parameter는 default로 설정하는 것이 좋음)
- ④ 입력정보를 모두 입력했으면 우측 하단의 BLAST 버튼 클릭
- ⑤ 분석이 완료되었으면 결과 파일을 다운로드 할 수 있습니다.

< 기타 분석 도구 화면 - BLAST >

VCF to MAF

- VCF to MAF는 Variant Call Format(VCF) 파일을 Minor Allele Frequency(MAF) 파일 형식으로 변환해주는 도구입니다.

① [파일 선택] 버튼을 통해 VCF 파일을 업로드 합니다.

(* Reference genome HG38 version을 사용하여 생성된 파일만 허용됩니다.)

② VCF 파일 업로드 후 [제출] 버튼을 누르면 변환이 시작됩니다.

③ 변환 후 결과 파일을 다운로드 할 수 있습니다.

< 기타 분석 도구 화면 - VCF to MAF >

< VCF to MAF 변환 결과 안내 메시지 >

Interpro

- Interpro는 알려진 단백질에서 발견되는 식별 가능한 기능을 새로운 단백질 서열에 적용할 수 있는 단백질 도메인, 단백질 패밀리 데이터베이스로써, 단백질 염기서열을 패밀리로 분류하고 도메인 및 중요 부위의 존재를 예측하여 기능 분석을 제공하는 리소스입니다.
- Interpro는 signatures 예측 모델을 사용합니다.
 - ① FASTA 형식의 서열정보를 입력 상자에 직접 입력하거나, [파일 선택] 버튼을 통해 아미노산 서열 정보를 담고 있는 FASTA 파일을 업로드 합니다.
 - ② 서열 정보 입력 또는 파일 업로드 후 [제출] 버튼을 누르면 분석이 시작됩니다.
 - ③ 분석 후 결과 파일을 다운로드 할 수 있습니다.

The screenshot shows the InterPro web interface. At the top, there are navigation links: '사이트 소개', '데이터 분석 예시', '사용자 데이터 분석' (highlighted), and '고객의 소리'. The main heading is 'Interpro'. Below it, there is a red circular icon with a white 'X'. A paragraph explains that InterPro tool is a protein sequence database that classifies sequences into domains and families and provides functional annotations. Another paragraph states that InterPro uses a machine learning model to predict protein functions and that it uses a signature-based approach. Below this, there is a form with two input fields: 'FASTA 형식의 서열' and '업로드 파일'. The 'FASTA 형식의 서열' field has a placeholder text '서열을 입력하세요'. The '업로드 파일' field has a '파일 선택' button and a '선택한 파일 없음' text. Below these fields is a blue '제출' button. At the bottom, there is a note: '참고: 단백질 서열은 fasta 포맷이어야 합니다.'

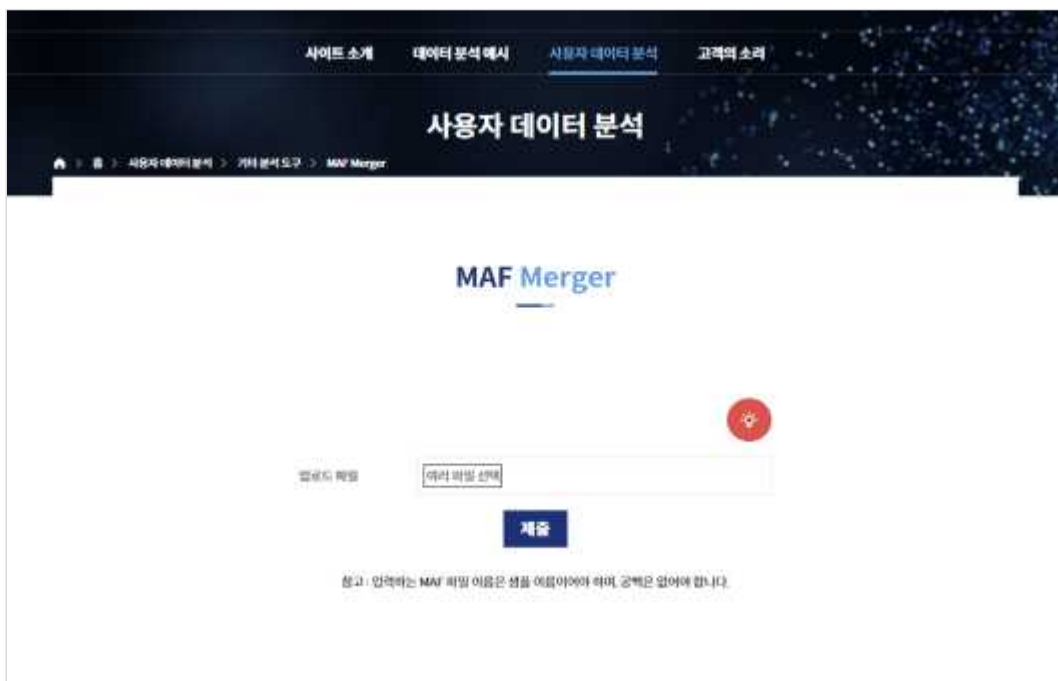
< 기타 분석 도구 화면 - Interpro >

The screenshot shows a message box with the text: '참고: 단백질 서열은 .fasta 포맷이어야 합니다.' and 'Your Results are ready kindly click link to download interpro-1700094723.708776'.

< Interpro 결과 안내 메시지 >

MAF Merger

- MAF Merger는 여러개의 MAF 파일을 입력해 1개의 파일로 병합하는 도구입니다.
 - 각 샘플 별 변이 정보를 담고 있는 여러개의 MAF 파일을 하나의 파일로 병합한 MAF 형식의 파일을 다운 받을 수 있고, 동시에 사용자 데이터 분석 서비스의 시각화 도구에 사용되는 DNA mutation 입력 파일로 활용할 수 있는 병합된 .tsv 파일도 다운 받을 수 있습니다.
- ① [여러 파일 선택] 버튼을 통해 병합시킬 MAF 파일을 업로드합니다.
 - ② 여러 개의 MAF 파일을 어볼드 후 [제출] 버튼을 누르면 병합이 시작됩니다.
 - ③ 병합 후 결과 파일을 다운로드 할 수 있습니다. (병합된 MAF 파일 및 DNA mutation 입력 파일)



< 기타 분석 도구 화면 - MAF Merger >



< MAF Merger 변환 결과 안내 메시지 >

RefVer Converter(Liftover)

- RefVer Converter(Liftover)는 사용자 VCF 파일의 참조 유전체 버전을 변환시켜주는 도구입니다.
 - 사용자 입력파일이 GRCh37(hg19) 버전이면 GRCh38(hg38)로 변환시켜주고, GRCh38(hg38) 버전의 입력파일을 넣으면 GRCh37(hg19) 버전으로 변환시켜주는 양방향 변환 도구입니다.
- ① [파일 선택] 버튼을 통해 변환할 VCF 파일을 업로드합니다.
 - ② 업로드 파일의 참조 유전체 버전을 선택합니다. (GRCh37 또는 GRCh38)
 - ③ 파일 업로드 및 업로드 파일의 참조 유전체 버전 선택 후 [제출] 버튼을 누르면 변환이 시작됩니다.
 - ④ 변환 후 결과 파일을 다운로드 할 수 있습니다.

< 기타 분석 도구 화면 - RefVer Converter(Liftover) >

참고 : VCF 파일의 GRCh37(hg19) 과 GRCh38(hg38) 버전 간의 양방향 변환 기능 지원
변환이 완료되었습니다. 다운로드 링크(vcf-converter-prod-1700094238-989676)를 통해 변환 결과들 다운로드하세요.

< RefVer Converter(Liftover) 변환 결과 안내 메시지 >

Dataframe Converter

- Dataframe Converter는 매트릭스 형태의 expression 파일을 Melted 형태로 변환시켜주는 도구입니다.
- 샘플에 대한 각 유전자 expression value를 Matrix 형태로 담고 첫 번째 행에 샘플에 대한 Type을 입력하여 Dataframe Converter Tool에 업로드하면 Matrix 형태의 입력 데이터를 Melted 형태로 변환합니다.

. 업로드 데이터 포맷

- 첫 번째 행 : 각 샘플의 Type 정보 (Normal sample : N / Tumor sample : T)
- 두 번째 행 : 각 샘플 ID
- 세 번째 행 : 유전자 이름 및 각 유전자에 대한 샘플의 value

① [파일 선택] 버튼을 통해 변환할 matrix 파일을 업로드합니다.

② 파일 업로드 후 [제출] 버튼을 누르면 변환이 시작됩니다.

③ 변환 후 결과 파일을 다운로드 할 수 있습니다.

	A	B	C	D	E	F	G	H	I	J	K
1	type	N	N	N	N	N	T	T	T	T	T
2	sample_id	sample_1	sample_2	sample_3	sample_4	sample_5	sample_6	sample_7	sample_8	sample_9	sample_10
3	5S_rRNA	0		0				0		0	0
4	7SK	0		0		0		0		0	
5	A1BG	24.74		49.28				13.34		5.77	12.46
6	A1BG-AS1	8.4	2.73	9.62	3.23	9.23					
7	A1CF		0			0	0.01		0	0.02	
8	A2M		265.98			199.2	576.11		174.55	112.07	
9	A2M-AS1	1.94			1.89	1.51		1.75	0.9		
10	A2ML1		1.09	0.02			0.83			24.44	0.53
11	A2ML1-AS1		0	0	0		0				0
12	A2ML1-AS2			0			0	0		0	0
13	A2MP1	0.37		0.31				0.04		0.16	0.28

< 변환 전 데이터 포맷 >

< 기타 분석 도구 화면 - Dataframe Converter >