

JU SeqWorkbench Alpha 사용자 가이드 — 한국어 초안

다른 언어: 한국어 | English

이 문서는 현재 구현된 알파 동작을 처음 사용하는 테스터가 재현 가능한 순서로 설명합니다. 메뉴명은 한국어 UI를 우선 사용하고 필요한 경우 현재 영문 라벨을 함께 적었습니다.

1. JU SeqWorkbench Alpha 개요

JU SeqWorkbench Alpha는 **Sanger/FASTA/MSA 서열 검토, 편집, 그룹화, 포인트 및 범위 중심 시각화**에 초점을 둔 로컬 데스크톱 애플리케이션입니다.

이 애플리케이션은 다음 도구가 아닙니다.

- NGS mapper
- variant-calling pipeline
- genome assembler
- clinical diagnostic system

알파 결과는 대표 데이터로 다시 확인하고 중요한 원본의 백업을 유지하십시오.

2. 지원 워크플로와 범위

1. FASTA/TXT 또는 Sanger AB1/ABI 파일을 열거나 현재 뷰어에 가져옵니다.
2. 서열 ID와 현재 작업 서열을 검토하고 필요한 부분만 편집합니다.
3. 필요하다면 사용자가 별도 설치한 MAFFT 또는 Clustal Omega로 정렬합니다.
4. 포인트 시각화, 범위 시각화, ID/이름 그룹화, AA 마커 분류, 유사도 군집화를 실행합니다.
5. 결과표·그림·FASTA를 내보내거나 작업 상태를 `.dvproj` 프로젝트로 저장합니다.

예상 출력은 작업에 따라 새 뷰어, 결과표, PNG/JPEG/SVG/PDF 그림, CSV, FASTA 또는 프로젝트 파일입니다. 계통수와 주석 레이어는 도움말의 추후 제공 항목이며 현재 분석 메뉴에서 실행할 수 없습니다.

3. 메인 창 구성

1. 상단 메뉴에서 파일, 편집, 보기, 정렬, 분석, 웹 리소스, 도움말을 선택합니다.

2. 도구막대에서 View/Edit , OVR/INS , InsV , ColType , DupV , ID , 좌표 간격, Dot , 잔기 색상을 조정합니다.
3. 왼쪽 서열 ID 패널에서 행을 선택합니다.
4. 오른쪽 눈금자와 서열 본문에서 위치와 문자를 검토합니다.

서열 관리 (추후 제공) 메뉴는 보이지만 비활성 상태입니다. ID 패널은 보기 > ID 패널 표시 또는 도구막대 ID 로 숨기거나 다시 표시합니다. 잔기 색상은 큰 정렬에서 느껴질 수 있습니다.

언어는 도움말 > 언어 에서 English 또는 한국어를 선택합니다. 설정은 로컬 QSettings에 저장되며 안내대로 애플리케이션을 다시 시작해야 전체 UI에 적용됩니다.

User Guide - Alpha

메인 뷰어 구성

핵심 영역

Menu: 파일·보기·분석·도움말 기능

Tool bar: 편집 상태와 빠른 실행

ID panel: 군주/서열 ID 선택

Seq panel: 염기 또는 아미노산 서열 표시

주요 토크

View/Edit, INS/OVR, Dot, Residue Colors, Coordinates, ID panel 표시, Clear all highlights는 화면 반응과 편집 대상에 직접 영향을 줍니다.

Tip: 메뉴 중 아직 알파 범위 밖이거나 제한적인 기능은 비활성화 또는 Planned 표시로 구분하는 것이 사용자 혼란을 줄입니다.

JU SeqWorkbench Alpha 02

4. 프로젝트 만들기과 열기

1. 파일 > 새로 만들기 또는 Ctrl+N 으로 빈 작업공간을 엽니다.
2. 기존 프로젝트는 파일 > 프로젝트 열기... 를 사용합니다.
3. 같은 프로젝트가 이미 열려 있으면 기존 창 활성화, 독립 사본 열기, 취소 중 하나를 선택합니다.
4. 독립 프로젝트 사본은 원본 경로에 직접 저장하지 않고 프로젝트 다른 이름으로 저장... 을 사용합니다.

프로젝트 파일은 .dvproj 가 기본이며 .dvproj.json 과 서명이 맞는 레거시 .json 을 열 수 있습니다. 일반 서열 열기 는 별도 뷰어를 만들고, 가져오기 는 현재 뷰어에 레코드를 추가합니다.

5. FASTA와 서열 데이터 가져오기

1. 새 창으로 보려면 **파일 > 열기 / Ctrl+O** 를 선택합니다.
2. 현재 작업에 추가하려면 **파일 > 가져오기...** / **Ctrl+I** 를 선택합니다.
3. 클립보드 FASTA는 **Ctrl+Alt+V** 로 새 레코드로 붙여넣습니다.
4. 중복 ID가 조정되었다는 안내가 나오면 새 접미사를 확인합니다.

FASTA, 일반 서열 TXT, ABI/ABI가 지원 경로에 포함됩니다. 한 줄 FASTA는 **>Sample_1 SEQUENCE** 처럼 ID와 서열을 구분할 수 있어야 합니다. 원시 서열의 종류가 모호하면 NT 또는 AA를 선택하라는 창이 나타날 수 있습니다.

6. 서열 검토와 편집

1. 도구막대에서 **Edit** 모드로 전환합니다.
2. 본문을 클릭해 커서를 놓거나 드래그해 범위를 선택합니다.
3. **Insert** 로 **INS** (삽입 모드)와 **OVR** (덮어쓰기 모드)을 전환합니다.
4. 문자 입력, **Delete / Backspace** , **Ctrl+X** , **Ctrl+V** 를 사용합니다.
5. **Ctrl+Z** / **Ctrl+Y** 로 결과를 검토합니다.

일반 드래그는 이전 본문 하이라이트를 대체하며, **Ctrl** 을 누른 채 드래그하면 새 하이라이트 레이어를 누적합니다.

User Guide - Alpha

하이라이트와 선택 상호작용

1. Column selection (1-10)

2. Row selection (1-10)

3. Ctrl+Drag action for cumulative highlighting

기본 동작

- 룰러 드래그: 컬럼 하이라이트
- 서열 영역 드래그: 가로 하이라이트
- Ctrl+Drag: 누적 하이라이트
- *Clear all highlights: 모든 하이라이트 제거

알파 제한사항

현재 일부 드래그 선택은 의도한 균주/행 경계를 넘어 더 넓게 지정될 수 있습니다. 여러 하이라이트가 있을 때 붙여넣기 대상은 가장 최근 활성 하이라이트입니다.

ID 패널과 서열 패널은 서로 다른 선택 모델을 사용하므로 여백 클릭, 선택 해제, 우클릭 메뉴의 동작이 동일하지 않을 수 있습니다.

JU SeqWorkbench Alpha

03

현재 행의 NT/RNA/AA 종류에서 허용되지 않는 문자는 차단됩니다. 하이라이트가 있을 때 **Delete** 는 가장 최근 하이라이트의 실제 서열 데이터를 삭제하고 **Ctrl+Delete** 는 모든 활성 하이라이트의 데이터를 삭제합니다. 표시만 지우려면 **모든 하이라이트 제거** 를 사용하십시오.

한 서열을 별도 창에서 편집하려면 ID를 오른쪽 클릭하고 **Open Sequence in Editor...** 를 선택합니다. 이 창은 모델리스이며 부모 뷰어와 Alt+Tab으로 전환할 수 있습니다. **Apply** 는 한 번의 되돌릴 수 있는 부모 변경을 만듭니다.

7. 서열 ID 편집

1. ID 패널에서 정확히 한 행을 선택합니다.
2. **Copy ID** 로 복사한 ID 또는 외부 프로그램의 적합한 일반 텍스트를 클립보드에 준비합니다.
3. ID 패널에 포커스를 둔 채 **Ctrl+V** 를 누르거나 오른쪽 클릭 메뉴의 ID 이름 변경 항목을 사용합니다.
4. 중복 이름에 붙은 **_2** , **_3** 접미사를 확인합니다.

오른쪽 클릭 메뉴의 주요 동작은 다음과 같습니다.

- **Copy ID** : 선택한 ID 이름을 복사합니다. 여러 행이 선택되면 선택한 ID들을 복사합니다.
- **Copy Sequence** : 정확히 한 strain의 현재 작업 서열만 헤더 없이 복사합니다. 이 클립보드 데이터는 ID 이름 변경에 사용할 수 없습니다.
- **Copy as FASTA** : ID와 서열을 FASTA 형식으로 복사합니다. 여러 행을 선택하면 여러 FASTA 레코드가 유지됩니다.
- **Open Sequence in Editor...** : 선택한 단일 strain을 별도의 모델리스 편집기에서 엽니다.
- **Paste ID / Rename ID from Clipboard** : 단일 ID를 **Copy ID** 데이터 또는 적합한 일반 텍스트로 변경합니다. **Copy Sequence** 에서 온 데이터는 ID로 적용되지 않습니다.
- **Delete Strain...** / **Delete Selected Strains...** : 선택한 strain을 삭제합니다. **Ctrl+Z** / **Ctrl+Y** 로 실행 취소/다시 실행할 수 있습니다.
- **Move Sequence...** : 정확히 한 strain을 선택해 시작한 뒤 destination ID를 클릭하면 source strain이 destination 바로 위로 이동합니다. **Esc** 로 대기 중인 이동을 취소할 수 있으며 이동은 **Ctrl+Z** / **Ctrl+Y** 를 지원합니다.
- **Move to Bottom** : 선택한 단일 strain을 목록 마지막으로 이동합니다. 이 이동도 **Ctrl+Z** / **Ctrl+Y** 를 지원합니다.

여러 ID가 선택되었거나 선택이 없으면 붙여넣기 이름 변경은 실행되지 않습니다. 여러 선택에서는 ID/FASTA 복사 및 선택 strain 삭제를 사용할 수 있지만, **Copy Sequence** 는 실행하지 않고 **Copy as FASTA** 사용을 안내합니다. 편집기 열기, ID 이름 변경, 이동은 단일 행 정책을 따릅니다. ID 패널의 **Ctrl+Alt+V** 는 ID 이름 변경이 아니라 기존의 새 서열 붙여넣기입니다.

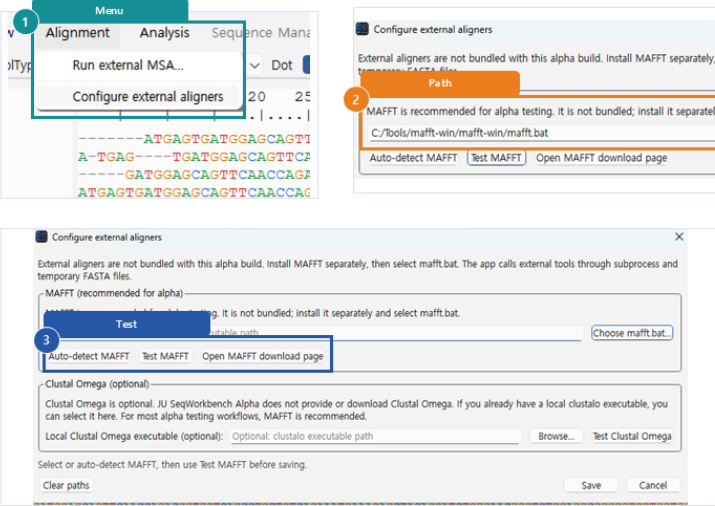
8. 외부 정렬 설정

1. MAFFT 또는 Clustal Omega를 애플리케이션과 별도로 설치합니다.
2. 정렬 > 외부 정렬 도구 설정 을 엽니다.
3. 로컬 실행 파일을 선택하거나 경로를 비워 둡니다.
4. 저장 후 설정 창을 닫습니다.

실행 파일은 JU SeqWorkbench에 포함되거나 자동 다운로드되지 않습니다. 로컬 사용자가 선택한 경로는 기존 QSettings에 저장됩니다. 경로가 비어 있으면 실행 시 PATH 에서 `mafft` 또는 `clustalo` 를 찾습니다. 설정 창 의 테스트 컨트롤은 선택한 로컬 실행 파일을 임시 FASTA로 확인합니다.

User Guide - Alpha

외부 MSA 도구 설정



1. Menu

2. Path

3. Test

설정 흐름

Alignment > Configure external aligners

MAFFT 또는 Clustal Omega 실행파일 경로 지정

Test로 실행 가능 여부 확인

*경로는 사용자 PC의 로컬 설정으로 저장

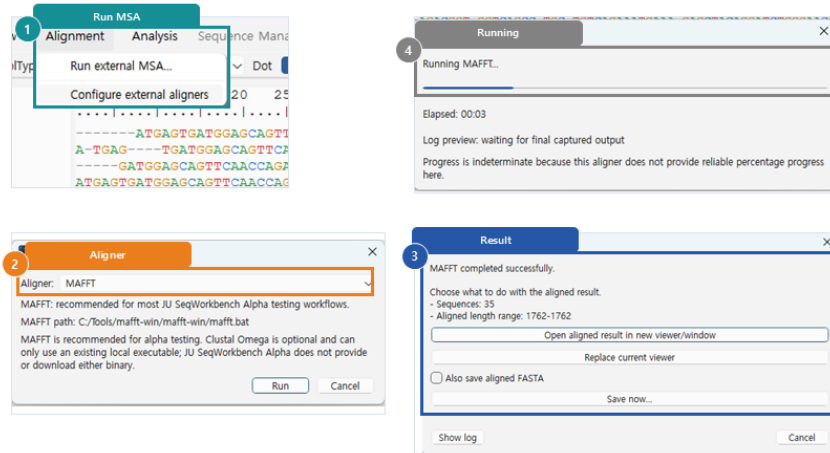
배포 정책

알파 배포본에는 MAFFT/Clustal Omega 바이너리가 포함되지 않습니다. 사용자가 별도로 설치한 실행파일을 연결하는 구조입니다.

9. MAFFT 또는 지원 정렬 도구 실행

1. 정렬할 현재 작업 서열이 두 개 이상인지 확인합니다.
2. 정렬 > 외부 MSA 실행... 을 선택합니다.
3. MAFFT 또는 Clustal Omega를 선택합니다.
4. 진행 창과 오류/로그를 확인합니다.
5. 성공 후 새 뷰어로 열기, 현재 뷰어 교체, 정렬 FASTA 저장 또는 결과 버리기 중 필요한 동작을 선택합니다.

MSA 실행과 결과 적용



결과 처리

MSA 실행 후 결과는 별도 결과 창으로 확인합니다. 현재 뷰어를 교체하거나 새 뷰어로 여는 경로를 명확히 선택할 수 있습니다.

주의

프로젝트 파일(.dvproj)은 직접 정렬 대상이 아닙니다. 앱은 현재 뷰어의 작업 서열을 임시 FASTA로 만들어 외부 도구에 전달합니다.

애플리케이션은 임시 FASTA를 만들고 외부 프로세스를 실행한 뒤 정렬 FASTA를 읽습니다. 현재 뷰어 교체는 편집 이력에 기록되며 하이라이트를 지웁니다. 실패 시 진단을 위해 임시 폴더가 유지될 수 있고 로그에서 경로와 명령을 확인할 수 있습니다.

10. NT, AA, DNA, RNA 변환

1. 변환할 ID 행을 선택합니다. 선택하지 않으면 모든 행이 대상입니다.
2. **Ctrl+D** / 아미노산 보기, **Ctrl+W** / DNA 보기, **Ctrl+E** / RNA 보기를 사용합니다.
3. 상태 표시줄의 변환·복구·건너뛰 수를 확인합니다.
4. 필요한 경우 **Ctrl+Z** 로 되돌립니다.

AA 보기는 NT 원본을 frame 0으로 번역합니다. DNA 보기는 보관된 NT 작업 상태로 돌아가며 AA 전용 레코드를 임의로 역번역하지 않습니다. RNA 보기는 T/U 표시를 전환합니다. 변환 시 본문/열 하이라이트가 지워지며 RNA 변환은 ID 선택도 지웁니다.

역방향, 상보, 역상보는 각각 **Ctrl+Alt+R**, **Ctrl+Alt+C**, **Ctrl+Alt+X** 입니다. 선택 영역이나 ID 행이 없으면 전체 서열에 적용됩니다.

11. 포인트 시각화

1. 분석 > 시각화 에서 AA, NT 또는 Codon 포인트 시각화를 선택합니다.

2. 위치를 1부터 시작하는 좌표로 입력합니다. 예: 10,25,50 .
3. 필터와 기준 서열을 선택하고 필요한 포함/제외 옵션을 설정합니다.
4. 실행 을 누릅니다.
5. Visualization, Counts, Detail, Warnings 탭을 검토합니다.

User Guide - Alpha

포인트 시각화: 조건 설정

Analysis

Sequence Management (Planned)

Web Resources

Visualization >

Clustering / Classification >

Point Visualization (AA)

Point Visualization (NT)

Point Visualization (Codon)

Region Visualization

입력 패널

Preset, Unit, Position, Stop codon handling, Filter, Reference를 설정합니다. AA/NT/Codon 모드에 따라 입력 좌표의 의미가 달라집니다.

출력 패널

같은 위치 집합으로 로고 플롯, 변이맵, 엔트로피, major allele frequency, 조성 막대 등 여러 형태의 결과를 전환해 볼 수 있습니다.

1

Input & Options

Preset: (AA) example_set1 Save Overwrite selected preset Delete

Unit: AA

Positions: 101,164,170,324,370,386,413,426,435,440,476

Stop codon handling: Truncate at first stop

☐ Include reference sequence
☐ Create per-strain detail table
☒ Trim length to a multiple of 3

Filter: (Header filter (subsetting))

Reference: (Automatic selection)

☐ Convert viewer display to the analysis basis (link/highlight)

Run

Run / Run again Close

2

Output tabs

Visualization Counts Detail Warnings

Visualization Logo plot

☒ Bits (information)
 ☒ Include

Press Run to generate results.

JU SeqWorkbench Alpha

06

포인트 시각화는 선택 위치의 로고, 히트맵, 이진/범주형 변이 지도, entropy, major allele frequency, 제외 수, 변이 막대, 구성 누적 막대 등을 제공합니다.

Counts/Detail 표의 행을 선택하면 표시 종류가 호환될 때 부모 뷰어 위치와 서열 행이 하이라이트됩니다. CSV와 그림 저장 버튼으로 결과를 내보낼 수 있습니다.

포인트 시각화: 표와 뷰어 하이라이트 연동

1 Result table

Visualization Counts Detail Warnings

View: Group by variant (IDs in one cell) Filter: Press Ctrl+F to focus; type to filter Clear Sort: Input order

N	Position	RefAA	MutAA	SeqIDs
0	101	G	*	
1	101	G	A	
2	101	G	D	
3	101	G	H	
4	101	G	I	
5	101	G	L	
6	101	G	M	
7	101	G	N	

3 Position

Sample_10, Sample_14, Sample_23, Sample_24, Sample_35, Sample_8

2 Viewer highlight

75 80 85 90 95 100 105 110 115 120

Sample_1
Sample_2
Sample_3
Sample_4
Sample_5
Sample_6
Sample_7
Sample_8
Sample_9
Sample_10
Sample_11
Sample_12
Sample_13
Sample_14
Sample_15
Sample_16
Sample_17
Sample_18
Sample_19
Sample_20
Sample_21
Sample_22
Sample_23
Sample_24
Sample_25
Sample_26
Sample_27
Sample_28
Sample_29
Sample_30
Sample_31
Sample_32
Sample_33
Sample_34
Sample_35

TOTFKYARK*KL*KSCK*YG*NCS*RRHFR*YSTNONTLVIG*CKCL
I*IQKVKIIEEWL*IIMIKLQUTETWL*MIHMKL*HLGHNLQMGLGE
NMPESENYRRVYNNLKTAYNGNALDDTHAQIVTFWLSVDANAGW
MPSENYRRVYNNLKTAYNGNALDDTHAQIVTFWLSVDANAGW
VHLMPESENYRRVYNNLKTAYNGNALDDTHAQIVTFWLSVDANAG
NMPESENYRRVYNNLKTAYNGNALDDTHAQIVTFWLSVDANAGW
YARK*KL*KSCK*FG*NCS*RRHFR*YSTNONTLVIG*CKCLGSLV
HLMPSESENYRRVYNNLKTAYNGNALDDTHAQIVTFWLSVDANAG
IQKVKIIEEWL*IIMIKLQUTETWL*MIHMKL*HLGHNLQMGLGE
NMPESENYRRVYNNLKTAYNGNALDDTHAQIVTFWLSVDANAGW
OQKVKIIEEWL*IIMIKLQUTETWL*MIHMKL*HLGHNLQMGLGE
IQKVKIIEEWL*IIMIKLQUTETWL*MIHMKL*HLGHNLQMGLGE
VI*IQKVKIIEEWL*IIMIKLQUTETWL*MIHMKL*HLGHNLQMGLGE
NMPESENYRRVYNNLKTAYNGNALDDTHAQIVTFWLSVDANAGW
NMPESENYRRVYNNLKTAYNGNALDDTHAQIVTFWLSVDANAGW
RK*KL*KSCK*FG*NCS*RRHFR*YSTNONTLVIG*CKCLGSLV*SR
K*KL*KSCK*FG*NCS*RRHFR*YSTNONTLVIG*CKCLGSLV*SR
YARK*KL*KSCK*FG*NCS*RRHFR*YSTNONTLVIG*CKCLGSLV
OQKVKIIEEWL*IIMIKLQUTETWL*MIHMKL*HLGHNLQMGLGE
IQKVKIIEEWL*IIMIKLQUTETWL*MIHMKL*HLGHNLQMGLGE
NMPESENYRRVYNNLKTAYNGNALDDTHAQIVTFWLSVDANAGW
NMPESENYRRVYNNLKTAYNGNALDDTHAQIVTFWLSVDANAGW
YARK*KL*KSCK*FG*NCS*RRHFR*YSTNONTLVIG*CKCLGSLV
I*IQKVKIIEEWL*IIMIKLQUTETWL*MIHMKL*HLGHNLQMGLGE
NMPESENYRRVYNNLKTAYNGNALDDTHAQIVTFWLSVDANAGW
NMPESENYRRVYNNLKTAYNGNALDDTHAQIVTFWLSVDANAGW
YARK*KL*KSCK*FG*NCS*RRHFR*YSTNONTLVIG*CKCLGSLV
YARK*KL*KSCK*FG*NCS*RRHFR*YSTNONTLVIG*CKCLGSLV
NMPESENYRRVYNNLKTAYNGNALDDTHAQIVTFWLSVDANAGW

결과 테이블

Counts 탭은 위치별 요약, Detail 탭은 군주변이 단위 상세 데이터를 제공합니다. 검색과 정렬로 관심 항목을 빠르게 좁힐 수 있습니다.

핵심 연결

테이블에서 선택한 변이 정보는 원본 뷰어의 해당 군주와 위치를 하이라이트해, 표와 정렬 화면 사이의 왕복을 줄입니다.

프리셋은 사용자 설정 위치에 저장되며 현재 배포 기본값에는 특정 데이터셋 위치가 없습니다.

12. 코돈 로고플롯

1. AA/NT 포인트 시각화에서 **Logo plot** 을 선택하거나 Codon 포인트 시각화에서 **Codon logo plot** 을 선택합니다.
2. gap, stop, unknown 포함 옵션과 bits/frequency 표시를 조정합니다.
3. 글꼴, 크기, 축 표시, publication 스타일을 선택합니다.
4. 다시 렌더링하고 그림을 저장합니다.

코돈 로고는 관찰된 완전한 세 문자 토큰을 그대로 씁습니다. 예를 들어 **GAT**, **-AT**, **A-T**, **AT-**, **NAA**, **TAA** 는 서로 다른 토큰입니다. **갭 포함 코돈**, **모호성 염기를 포함한 코돈**, **종결 코돈** 분류는 색상과 범례에만 사용되며 분류 명이 로고 글리프로 표시되지 않습니다. 표준 코돈 색은 코돈 정체성별 색입니다.

포인트 시각화: 그림 결과 조정

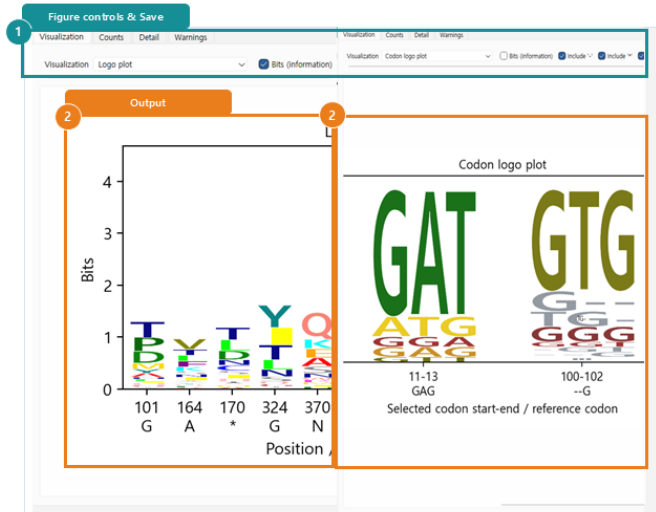


그림 옵션

Title, 축 라벨, 눈금, 스케일, 텍스트 크기, 폰트, 스타일을 조정한 뒤 Rerender로 다시 그릴 수 있습니다.

시각화 영역

Logo plot 외 여러 시각화 기능이 있습니다. 시각화 기능 및 염기의 종류에 따라 시각화 방식이 다를 수 있습니다. (Codon의 Logo plot의 경우 3글자 단위로 시각화)

저장

Save figure로 PNG/JPG 형태의 그림을 저장합니다. 논문·보고서용 이미지는 이후 PPT에서 번호와 캡션을 붙여 정리할 수 있습니다.

13. 범위 시각화

1. 분석 > 시각화 > 범위 시각화 를 엽니다.
2. 하나 이상의 시작-끝 범위를 입력합니다.
3. 기준 서열과 프로파일/요약 옵션을 선택합니다.
4. 실행 후 그림과 범위 지표 표를 검토합니다.
5. Export CSV 또는 Save figure 로 저장합니다.

범위 시각화: 구간 입력과 그림



구간 정의

Preset 또는 직접 입력으로 여러 region을 지정합니다.
Panel, Concatenate 등 레이아웃을 선택할 수 있습니다.

Metric

Entropy, gap/unknown rate, non-reference rate, mutation map, strain-wise profile 등 구간 중심의 지표를 계산합니다.

성능 메모

Lazy rendering은 아직 구현되지 않았습니다. 많은 서열 또는 색상 표시가 켜진 상태에서는 렌더링이 느릴 수 있으며, 필요 시 Residue Colors를 끄면 반응성이 개선됩니다.

범위 시각화는 범위별 변이, entropy, 프로파일과 부담 요약을 다룹니다. 그림은 PNG/JPEG와 SVG/PDF 경로를 지원합니다. 현재 창 내부는 영어 우선이며, 결과 셀/그림 점을 클릭해 부모 뷰어를 하이라이트하는 연동은 구현되지 않았습니다.

범위 시각화: Metrics table

region_label	unit	position	entropy	observed_tokens	token_summary	valid_count	excluded_count
0 Region A	aa	10	2.5299789188029957	A,G,L,P,Q,R,T,V,W	A:1; G:14; L:1; P:2; Q:2; R:4; T:2; V:7; W:1	34	1
1 Region A	aa	11	2.5327195351219203	C,D,G,R,R,S,T,V,W	C:1; D:2; G:13; P:1; R:3; S:3; T:1; V:5; W:1	30	5
2 Region A	aa	12	2.5308006218550014	A,C,G,L,Q,S,T,V,W	A:1; C:1; G:3; L:2; Q:13; S:6; T:1; V:1; W:3	31	4
3 Region A	aa	13	2.6652399286796387	A,C,G,L,M,P,Q,R,S,V	A:1; C:1; G:2; L:7; M:1; P:13; Q:2; R:1; S:3; V:2	33	2
4 Region A	aa	14	2.70768021233142	A,C,G,K,L,Q,R,S,V,W	A:14; C:2; G:1; K:2; L:5; Q:2; R:2; S:4; V:1; W:1	34	1
5 Region A	aa	15	2.8501169519558163	A,C,E,L,N,P,Q,R,S,V,W	A:1; C:5; E:3; L:1; N:1; P:2; Q:1; R:1; S:7; V:11; W:1	34	1
6 Region A	aa	16	2.9220024616785656	A,C,E,K,L,M,N,Q,R,S	A:3; C:4; E:7; K:2; L:2; M:2; N:1; Q:1; R:10; S:2	34	1
7 Region A	aa	17	3.154559938249354	C,E,I,K,M,N,Q,R,S,T,V	C:2; E:1; I:2; K:4; M:5; N:8; Q:4; R:1; S:3; T:1; V:2	33	2
8 Region A	aa	18	2.496877994577654	A,C,D,E,K,L,M,Q,R,Y	A:1; C:2; D:1; E:13; K:10; L:1; M:1; Q:1; R:3; Y:1	34	1
9 Region A	aa	19	3.0130059710789405	A,C,E,K,L,M,N,Q,R,S,T	A:1; C:1; E:5; K:4; L:3; M:2; N:2; Q:1; R:10; S:2; T:1	32	3
10 Region A	aa	20	3.0520121205143798	A,E,G,I,K,L,Q,R,S,T,Y	A:9; E:4; G:2; I:1; K:5; L:5; Q:2; R:1; S:1; T:1; Y:2	33	2
11 Region A	aa	21	3.0487949406953985	A,D,E,G,K,L,Q,R,S,T,W,Y	A:1; D:1; E:1; G:1; K:2; L:2; Q:5; R:5; S:2; T:10; W:1; Y:1	32	3
12 Region A	aa	22	3.124226735509437	A,D,E,G,I,K,L,Q,R,S,T,Y	A:3; D:4; E:1; G:10; I:3; K:1; L:2; Q:2; R:1; S:2; T:1; Y:1	31	4
13 Region A	aa	23	3.136324711025711	D,G,I,K,L,N,Q,R,S,T,W,Y	D:2; G:3; I:1; K:1; L:5; N:1; Q:1; R:2; S:10; T:4; W:3; Y:1	34	1
14 Region A	aa	24	2.468090781663233	D,E,G,I,L,N,S,T,V,W,Y	D:1; E:4; G:18; I:1; L:2; N:1; S:2; T:2; V:1; W:1; Y:1	34	1
15 Region A	aa	25	2.9357417972482365	A,E,G,I,L,N,R,S,T,V	A:1; E:1; G:6; L:1; N:9; R:3; S:5; T:4; W:3; Y:1	34	1
16 Region A	aa	26	2.147805973542275	E,G,I,L,N,R,S,T,V	E:1; G:20; I:1; L:2; N:1; R:3; S:1; T:2; V:3	34	1

표 보기

Show metrics table은 전체 구간, 구간별 탭, long-form detail 등 계산 결과를 표로 확인하는 창입니다.

정렬과 검색

Table view, Sort column, Direction, Scope, Search 옵션으로 필요한 행을 좁혀볼 수 있습니다.

내보내기

CSV Export로 현재 표를 저장합니다. 매우 큰 detail table은 미리보기와 전체 export 범위가 다를 수 있습니다.

14. ID 기반 그룹화

User Guide - Alpha

그룹화와 분류 기능 개요



Similarity clustering

NT 또는 AA 서열의 유사도 기준으로 균주를 묶고, 결과 표와 그룹별 FASTA를 생성합니다.

AA Marker classification

특정 AA 위치의 조건 조합으로 그룹을 정의합니다.

예: 10번 위치 A, 25번 위치 G.

ID/name grouping

서열 ID 또는 description에 포함된 키워드를 기반으로 그룹을 분류합니다.

예: Group_A, Year_2024.

공통 원칙

분류 성공 시 설정창은 닫히고, 결과 테이블이 앞으로 표시됩니다. 오류나 검증 실패가 있으면 설정창을 유지해 수정할 수 있습니다.

각 분류 기능은 현재 뷰어의 작업본을 기준으로 실행되며, 필요한 경우 그룹별 FASTA를 내보낼 수 있습니다.

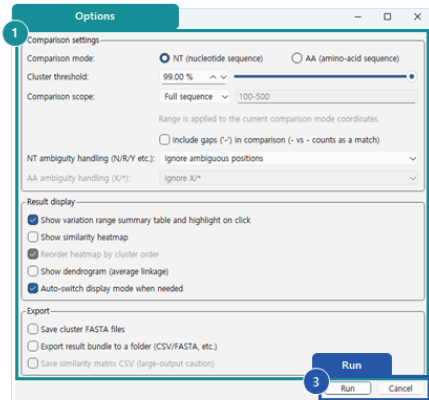
JU SeqWorkbench Alpha

11

1. 분석 > 군집화 / 분류 > ID/이름 기반 그룹화 를 엽니다.
2. 새로 만들기 로 규칙셋을 만들고 그룹 규칙을 추가합니다.
3. Group_A 같은 독립적인 그룹 이름과 한 줄당 하나의 ID 패턴을 입력합니다.
4. 대소문자, 경계 일치, 다중 일치, 표시 형식 옵션을 검토하고 저장합니다.
5. 결과표 표시 및 그룹 FASTA 저장 여부를 선택한 뒤 실행합니다.

현재 뷰어의 서열 ID/이름 문자열을 규칙 패턴과 비교해 그룹을 만듭니다. 성공하면 설정 창이 닫히고 완료 안내 뒤 결과표가 앞으로 표시됩니다. 검증 또는 실행 실패 시 창은 열린 상태로 유지됩니다. 결과표의 일반 행 선택은 부모 뷰어 하이라이트와 연결되지 않습니다.

Similarity clustering

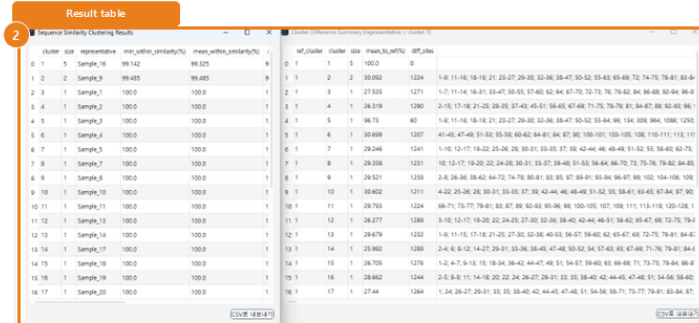


옵션

NT/AA 기준, 유사도 임계값, gap/unknown 처리, 비교 범위를 지정합니다.

결과

클러스터 요약, pairwise similarity, 다른 구간 표시, heatmap 또는 FASTA bundle export를 확인할 수 있습니다.

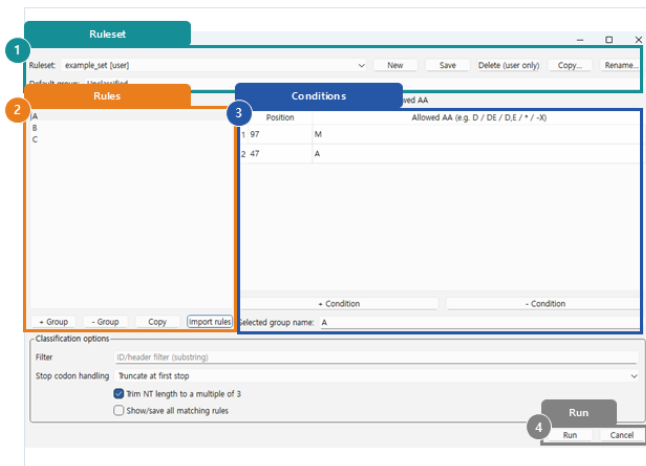


15. AA 마커 그룹화

1. 분석 > 군집화 / 분류 > AA 마커 기반 분류 를 엽니다.
2. 기본 내장 규칙셋이 없으므로 새로 만들기 로 사용자 규칙셋을 만듭니다.
3. 그룹과 조건을 추가하고 1부터 시작하는 Position과 Allowed AA를 입력합니다.
4. 규칙셋을 저장하고 stop 처리, 다중 일치, 결과표/그룹 FASTA 옵션을 선택합니다.
5. 실행 을 누르고 결과를 검토합니다.

Position은 숫자만 입력되며 기존 범위 검증을 거칩니다. Allowed AA는 대문자로 정규화되고 지원 문자만 입력됩니다. 불완전하거나 잘못된 편집 버퍼는 모델에 저장되지 않습니다. 다른 규칙으로 이동할 때 경고 후 미적용 값을 버리고 클릭한 규칙을 표시합니다. 완전히 빈 새 조건은 경고 없이 버릴 수 있습니다.

AA Marker classification



Ruleset

독립된 마커 규칙 묶음입니다. 사용자 정의 ruleset을 저장·복사·삭제할 수 있습니다.

Rule과 Condition

각 Rule은 하나의 그룹을 의미하고, 여러 condition을 조합해 특정 AA 위치의 허용 문자를 지정합니다.

입력 검증

Position은 숫자만 허용됩니다. Allowed AA는 유효한 아미노산 문자만 입력되며, 불완전 조건은 적용되지 않습니다.

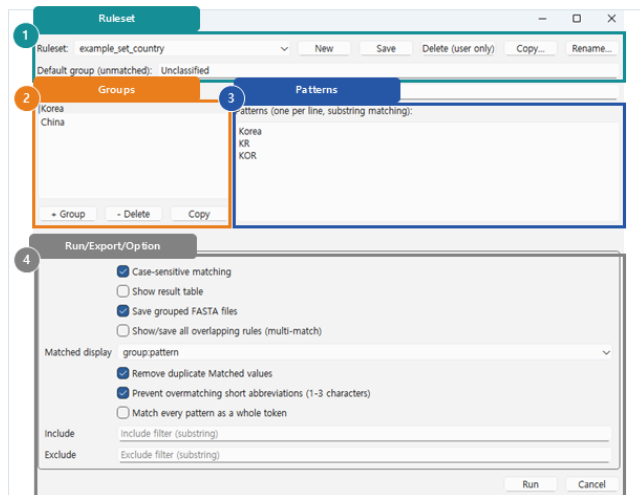
성공 시 설정 창을 닫고 안내 뒤 결과표를 앞으로 표시합니다. 실패 시 창을 유지합니다. 사용자 규칙은 기존 `typing_rulesets.json` 저장 형식을 사용합니다.

16. 유사도 군집화

1. 분석 > 군집화 / 분류 > 유사도 기반 군집화 를 엽니다.
2. NT 또는 AA 비교 모드와 군집 임계값을 선택합니다.
3. gap 및 모호성 처리, 전체 서열/수동 범위, 히트맵·덴드로그램·변이 구간 표 옵션을 설정합니다.
4. 선택된 ID 행 또는 선택이 없을 때 전체 현재 서열을 대상으로 실행합니다.
5. 결과표와 선택적 그림/FASTA/결과 묶음을 검토합니다.

최소 두 서열이 필요합니다. 길이가 다르면 정렬 확인을 요청합니다. 수동 범위는 현재 비교 모드의 1-based 좌표입니다. 성공 시 설정 창이 닫히고 결과표가 앞으로 표시되며 실패 시 창을 유지합니다. 결과표 행은 군집 구성원을, 변이 구간 표 행은 가능한 경우 구성원과 열 범위를 부모 뷰어에 하이라이트합니다.

ID/name grouping



키워드 기반 분류

ID 또는 description에 포함된 단어-패턴을 기준으로 그룹을 지정합니다. 각 패턴은 한 줄에 하나씩 입력합니다.

매칭 옵션

대소문자 구분, 보수적 경계 매칭, first/all match, matched display 방식을 선택할 수 있습니다.

Export

분류 결과 테이블과 그룹별 FASTA export를 지원합니다. AA-only 레코드는 별도 *_aa_only FASTA로 분리됩니다.

17. 결과표와 뷰어 하이라이트

1. 포인트 Counts/Detail 또는 유사도 결과표에서 행을 선택합니다.
2. 부모 뷰어의 ID 행과 위치 하이라이트를 확인합니다.
3. 표시 모드 불일치 안내가 나오면 분석 모드와 NT/AA 표시를 확인합니다.
4. 필요하면 모든 하이라이트 제거 로 연결 결과를 지웁니다.

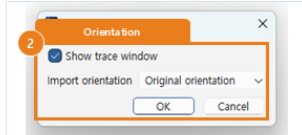
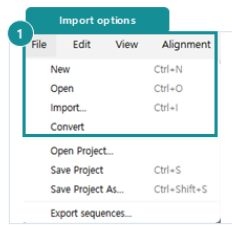
모든 결과표가 부모 뷰어와 연결되는 것은 아닙니다. 현재 확인된 연결은 포인트 시각화와 유사도 결과표입니다. 범위 시각화, AA 마커, ID/이름 그룹화 결과표의 일반 행 선택에는 같은 연결이 없습니다. 표에서 **Ctrl+C** 는 선택 셀을 헤더와 함께 TSV 형태로 복사합니다.

18. AB1 크로마토그램 워크플로

1. AB1/ABI를 새 문서로 열려면 파일 > 열기 를 사용합니다.
2. 현재 뷰어에 추가하려면 파일 > 가져오기... 에서 AB1을 선택합니다.
3. 트레이스 창 표시와 원본/역상보 방향을 선택합니다.
4. 트레이스, basecalled sequence, 품질 요약과 경고를 검토합니다.
5. 시작/끝 trim 범위를 조정하고 필요한 방향으로 가져옵니다.

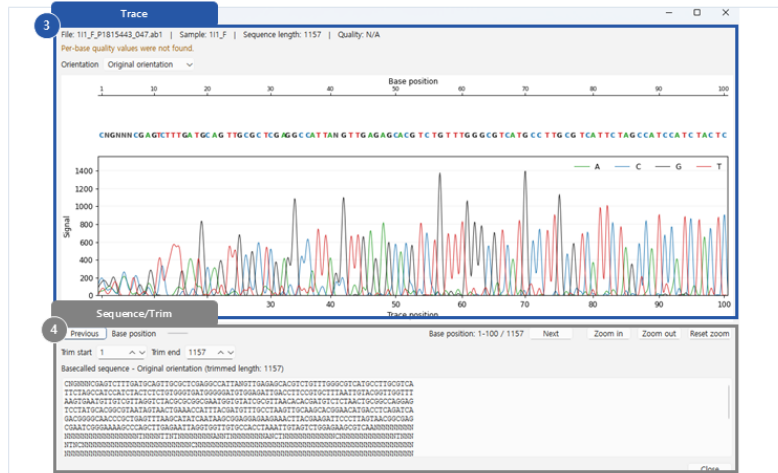
애플리케이션은 AB1에 이미 저장된 basecall을 읽으며 새 basecalling을 수행하지 않습니다. 트레이스 창은 위치 스크롤, 이전/다음, 확대/축소, 원본/역상보 표시를 제공합니다. 여러 AB1 가져오기에서 중복 ID는 조정됩니다.

Sanger AB1 chromatogram workflow



Trace preview

AB1/ABI 파일을 열어 chromatogram, basecalled sequence, quality summary를 확인합니다.



Import

Original 또는 Reverse complement 방향을 선택하고, 필요한 경우 trim 범위를 조정해 뷰어로 가져옵니다.

19. 표와 그림 내보내기

1. 메인 서열은 **파일 > 서열 내보내기...** 에서 FASTA, CSV 서열 표 또는 TXT로 저장합니다.
2. 결과표의 CSV 내보내기 버튼 또는 오른쪽 클릭/복사 기능을 사용합니다.
3. 포인트·범위·유사도 그림의 저장 버튼을 사용합니다.
4. 그룹화/분류의 그룹 FASTA 옵션을 선택하면 저장 대상과 덮어쓰기를 확인합니다.

프로젝트 저장과 데이터 내보내는 별개입니다. **Ctrl+S** 는 프로젝트 저장이며 FASTA 원본 덮어쓰기가 아닙니다. AA 전용 레코드는 AA FASTA로 분리될 수 있고 AA에서 NT로 임의 변환하지 않습니다. 복제 뷰어는 프로젝트 및 뷰어 수준 서열 데이터 내보내기를 차단합니다.

20. 프로젝트 저장과 다시 열기

1. **Ctrl+S** 또는 **파일 > 프로젝트 저장** 을 사용합니다.
2. 처음 저장하거나 독립 프로젝트 사본이면 저장 경로를 선택합니다.
3. 작업을 종료한 뒤 **프로젝트 열기...** 로 다시 엽니다.
4. 서열 순서, 현재 작업 상태와 표시 설정을 확인합니다.

프로젝트는 원본/현재 NT·AA 작업 상태와 관련 뷰어 상태를 보관합니다. 실행 취소/다시 실행 이력은 프로젝트에 저장되지 않으므로 다시 연 세션에서는 새 이력이 시작됩니다. 알파 테스트에서는 프로젝트 파일도 별도 백업하십시오.

21. 문제 해결

1. 시작이 늦으면 시작 화면이 먼저 나타나는지 확인하고 첫 메인 창 이벤트까지 기다립니다.
2. 정렬 실행 파일 오류가 나면 설정 경로와 `PATH` 를 확인합니다.
3. 유사도/포인트 하이라이트가 보이지 않으면 NT/AA 비교 모드와 현재 표시 종류를 맞춥니다.
4. FASTA 붙여넣기가 본문에서 거부되면 `Ctrl+Alt+V` 를 사용합니다.
5. 여러 레코드 복사는 `Copy Sequence` 가 아니라 `Copy as FASTA` 를 사용합니다.
6. 메인 `Ctrl+F` 가 검색하지 않는 것은 현재 알파의 알려진 제한입니다.
7. 대용량 분석 중에는 사전 확인과 작업 중 표시를 기다리고 결과 생성이 끝났는지 확인합니다.

오류 메시지가 있는 설정 창은 닫히지 않으므로 값을 수정해 다시 실행할 수 있습니다. 외부 MSA 실패 창에서는 `Show log` 로 명령, 출력, 임시 경로를 확인할 수 있습니다.

22. 알려진 제한

요약하면 외부 정렬 도구 별도 설치, 세션 간 Undo 미복원, 대용량 성능, 제한된 AB1 고급 기능, 메인 찾기 미구현, 일부 영어 우선 UI와 제한된 결과표 연동이 있습니다. 정확한 목록은 `known_limitations_ko.md`를 참조하십시오.

23. 피드백과 버그 보고

1. `도움말 > 버전 정보` 에서 현재 알파 버전을 확인합니다.
2. 재현에 사용한 입력 종류(FASTA, AB1, 프로젝트), 서열 수, 대략적인 길이를 기록합니다.
3. 클릭 순서, 기대 결과, 실제 결과와 오류 메시지를 기록합니다.
4. 개인·민감 서열을 제거한 중립 예제(`Sample_1` , `Sample_2` , 위치 `10` , `25` , `50`)로 재현 가능하면 함께 설명합니다.
5. `도움말 > 사용자 가이드` 에 표시된 현재 Issues 또는 Discussions 경로로 보고합니다.

프로젝트나 실제 서열을 첨부하기 전에 공개 가능한 데이터인지 확인하십시오. 로컬 실행 파일 경로와 사용자 프로필 경로는 보고서에서 가리십시오.