

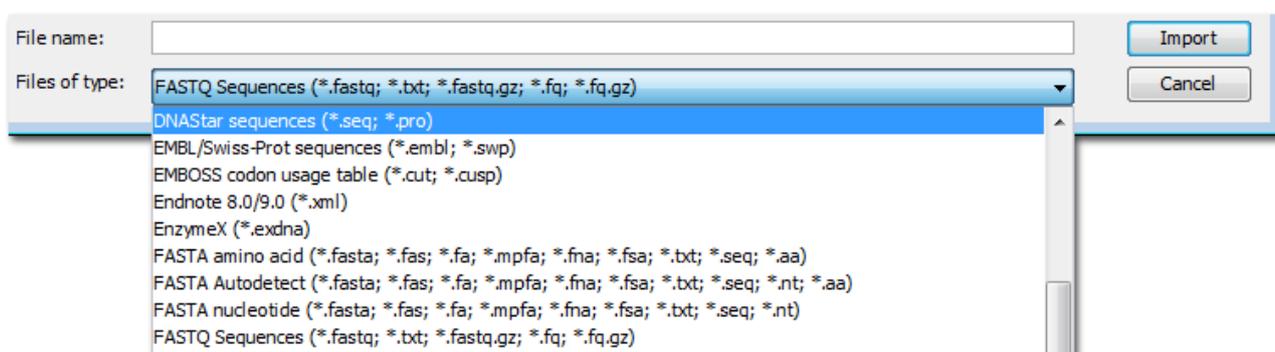
# Geneious Prime でシーケンス解析



## 第 1 回 シーケンスデータのインポート

Geneious Prime では、サンガーや NGS ショート/ロングリードシーケンサー、その他の解析ソフトウェアなどから生成される幅広いフォーマットのファイルをスマートにインポートすることができます。

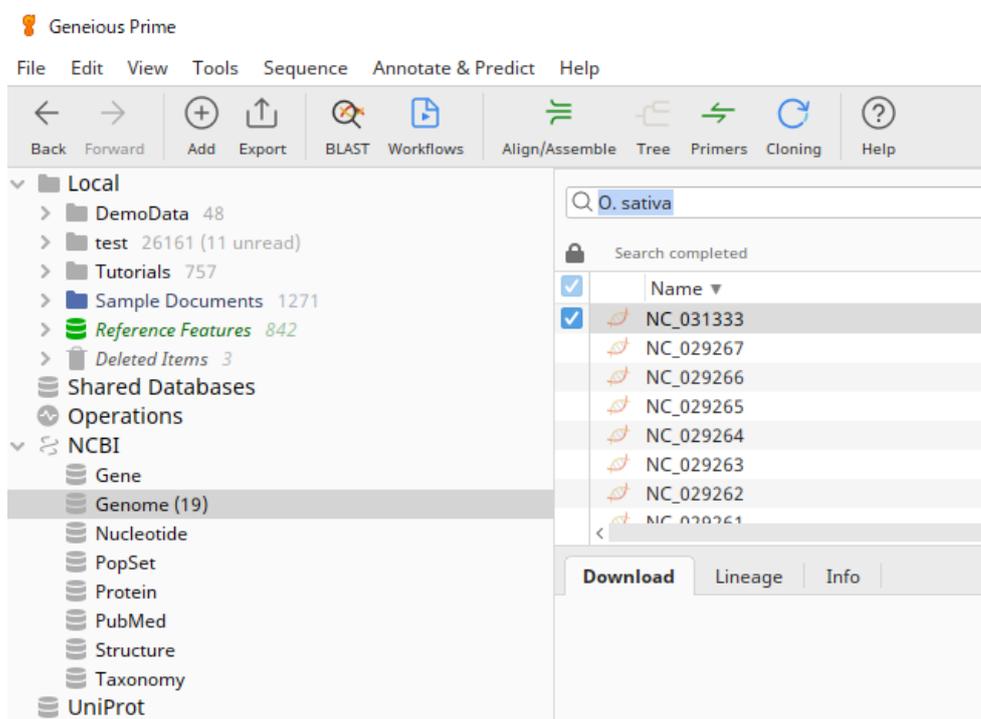
ほとんどの場合、ファイルやフォルダをハードディスクやネットワークドライブなどから Geneious Prime に直接ドラッグ&ドロップするだけで、自動的にフォーマットとフォルダ構造が認識されてインポートされます。もし自動認識がうまくいかない場合は、ツールバーの Add ボタンか File → Import → Files... から、ファイルのフォーマットを選択してインポートしてください。



SAM、BAM、GFF、BED、VCF、FASTA フォーマットのファイルを一括選択してドラッグ&ドロップすることで、シーケンス、アノテーション、アセンブリ情報を一度にインポートすることもできます。

また公共データベースである NCBI と UniProt からデータを直接インポートすることもできます。

画面左側ソールパネルの下側にある NCBI または UniProt フォルダを選択し、画面上側の検索ボックスに検索ワードを入れて Enter キーか Search ボタンをクリックすると検索がスタートします。



ドキュメントテーブルに表示された検索結果から、インポートしたいデータを選択し、ドキュメントビューにある Download ボタンをクリックしてダウンロードした後、ローカルデータベース内のフォルダにファイルをドラッグして移動してください。

## Geneious Prime にインポートできるデータフォーマットリスト

フォーマット	拡張子	データ形式	主な由来
BED	*.bed	Annotations	UCSC
Common Assembly Format	*.caf	Contigs	Sequencher
Clone Manager molecule	*.cm5	Sequences and annotations	Clone Manager
Clustal	*.aln	Alignments	ClustalX
CSFASTA	*.csfasta	Color space FASTA	ABI SOLiD
Comma/Tab Separated Values	*.csv, *.tsv	Spreadsheet files	Microsoft Excel
DNASTar	*.seq, *.pro	Nucleotide & protein sequences	DNASTar
DNA Strider	*.str	Sequences	DNA Strider (Mac program), ApE
Embl/UniProt	*.embl, *.swp	Sequences	Embl, UniProt
EMBOSS codon usage table	*.cusp, *.cut	Codon usage table	EMBOSS cusp tool
Endnote (8.0 or 9.0) XML	*.xml	Journal article references	Endnote, Journal article websites
FASTA	*.fasta, *.fas, *.fasta.gz etc.	Sequences, alignments	PAUP*, ClustalX, BLAST, FASTA
FASTQ	*.fastq, *.fq, *.fastq.gz etc.	Sequences with quality	Illumina and other NGS sequencers
GCG	*.seq	Sequences	GCG
GCG codon usage table	*.cod	Codon usage table	GCG CodonFrequency tool, <a href="https://www.kazusa.or.jp/codon/">https://www.kazusa.or.jp/codon/</a>
GenBank	*.gb, *.xml	Nucleotide & protein sequences	GenBank
Geneious	*.xml, *.geneious	Preferences, databases	Geneious
Geneious Education	*.tutorial.zip	Tutorial, assignment etc.	Geneious
GFF, GFF3, GTF	*.gff, *.gff3, *.gtf	Annotations	NCBI, Ensembl and other genome browsers
MEGA	*.meg	Alignments	MEGA
Molecular structure	*.pdb, *.mol, *.xyz, *.cml, *.gpr, *.hin, *.nwo	3D molecular structures	3D structure databases and programs
Newick	*.tre, *.tree, etc.	Phylogenetic trees	PHYLIP, Tree-Puzzle, PAUP*, ClustalX
Nexus	*.nxs, *.nex	Trees, Alignments	PAUP*, Mesquite, MrBayes & MacClade
PDB	*.pdb	3D Protein structures	SP3, SP2, SPARKS, Protein Data Bank
PDF	*.pdf	Documents, presentations	Adobe Writer, LATEX, Miktex
Phrap ACE	*.ace	Contig assemblies	Phrap/Consed
PileUp	*.msf	Alignments	pileup (gcg)
PIR/NBRF	*.pir	Sequences, alignments	NBRF PIR
Qual	*.qual	Quality file	Associated with a FASTA file
Raw sequence text	*.seq	Sequences	Any file that contains only a sequence
Rich Sequence Format	*.rsf	Sequences, alignments	GCGs NetFetch
SAM/BAM	*.sam, *.bam	Contigs	SAMtools
Sequence Chromatograms	*.ab1, *.scf	Raw sequencing trace & sequence	Sequencing machines
SnapGene sequence Text/html	*.dna, *.prot .txt, .rtf, .html	Sequences and annotations Any text	SnapGene Simple text editors
VCF	*.VCF	Annotations	1000 Genomes Project
Vector NTI sequence	*.gb, *.gp	Nucleotide & protein sequences	Vector NTI
Vector NTI/AlignX alignment	*.apr	Alignments	Vector NTI, AlignX
Vector NTI Archive	*.ma4, *.pa4, *.oa4, *.ea4, *.ca6	Nucleotide & protein sequences, enzyme sets and publications	Vector NTI
Vector NTI/ContigExpress	*.cep	Nucleotide sequence assemblies	Vector NTI
Vector NTI database	VNTI Database	Nucleotide & protein sequences, enzyme sets and publications	Vector NTI