

TOMY DIGITAL BIOLOGY CO., LTD.



第37回 自動解析ワークフローの作成2

(ステップの追加)

Geneious では、ワークフロー機能を使用することで、よく使用する解析の組み合わせを実行するために必要な別々のステップをグループ化し、解析を自動化することができます。

<u>チュートリアル用のデータがこちらからダウンロード可能</u>です。ダウンロードした zip ファイルは解凍せずに Geneious Prime にドラッグ&ドロップすることでインストールできます。

前回の記事で、新しいワークフローにステップを追加する準備ができましたので、Add step → Add operation で操作ステップを追加していきます。

• • •	Crea	te Workflow	
Workflow Name:	Sanger analysis		
Author:	Geneious Prime		
Description:	Sanger sequences trimming, de novo assembly, multiple alignment and tree building		
Icon:	🙏 Chromatogram	~	Choose Custom Icon
0	Share (read-only) with other Shared	Database users	
6	Share (read-only) with other Geneio	us Cloud users	
+ Add Step $-$	Delete Step 🖉 View/Edit Options 🤺	ackslash Move Down	⑦ Help
 Add Operation Add Recently Sor Each Door 	Image: Add Operation (from 111 available) Add Recently Used Operation Image: Ser Each Document		
용 For Each Seq 양 Group Docur 양 Group Seque	uence / Extract Sequences From List nents ences		
 Add docume Combine Wite Save Docume 	nt chosen when running workflow h Earlier Document(s) ents / Branch		
 ✓ Rename Document(s) □ Copy Property Between Documents Q Filter Documents ↓= Sort Documents 			
🖉 Custom Java	Code		
			Cancel OK

最初に実行したいのは、シークエンスの 3'と 5'末端の低品質な塩基をトリミングすることです。 Filter ボックスに trim と入力すると、利用可能な操作の候補が表示されますので、Trim Ends を 選択して OK します。



Trim Ends がワークフローに追加されます。

• 0 •	Create Workflow			
Workflow Name:	Sanger analysis			
Author:	Geneious Prime			
Description:	Sanger sequences trimming, de novo assembly, multiple alignment and tree building			
Icon:	🕅 Chromatogram			
0	Share (read-only) with other Shared Database users			
Share (read-only) with other Geneious Cloud users				
+ Add Step $-$	Delete Step $ \mathscr{O}$ View/Edit Options $ \uparrow $ Move Up $ \downarrow $ Move Down $$ $$ $$ $$ Help $$			
Trim E	Trim Ends			
Options: Trim regions				
	Cancel OK			

Trim Ends のオプションを確認し、変更するためには、Trim Ends のステップをダブルクリックします。今回の例では、Error Probability Limit を 0.01 に変更して OK します。 3'と 5'末端でエラー率が 1%を超える(QV20 以下の)低品質な塩基をトリミングする設定です。

	Edit Trim Ends				
Options to expose to user when workflow is run					
 Expose no options 					
C Expose all options					
Expose some options					
Optionally label exposed options as:	Access exposed options via button ?				
Expose: action	✓ With Alternative Label: (Optional) +				
All Operation Options (those not exposed to workfle	ow user and default values for options that are exposed)				
 Annotate new trimm Demose new trimm 	ned regions (regions will be excluded from assembly and consensus)				
Remove existing trim	nmed regions from sequences				
Trim vectors:	UniVec (High sensitivity) (will be automatically \sim $-$				
Minimum BLAST alignment score	e: 16 🗘				
Trim primers:					
	Choose				
	Allow Mismatches: 5				
	Minimum Match Length: 5				
Error Probability Limit:	✓ Error Probability Limit: 0.01 ♦ (decrease to trim more)				
Trim regions	Trim regions with more than a 4.3% chance of an error per base				
Maximum low quality bases:					
Maximum ambiguities:					
Trim 5' End	At least 0 0 bp				
V Trim 3' End	At least 0 🗇 bp				
Maximum length after trim:	1,000 🗘 (Trim excess from 3' end)				
More Options V	Cancel				

同じ手順で、2 つめのステップとしてデノボアセンブルを追加します。 Add step \rightarrow Add operation で、Align/Assemble \rightarrow De Novo Assemble を選択します。追加された De Novo Assemble ステップをダブルクリックし、下図のようにオプションを変更します。

Assemble reads (eg. Sanger or NGS) without using a reference Deptions to expose to user when workflow is run P Expose no options Expose all options Exposes some options Optionally label exposed options as: Expose: None Vith Alternative Label: (Optional) + - UI Operation Options (those not exposed to workflow user and default values for options that are exposed) Data Dissolve contigs and re-assemble Dissolve contigs and re-assemble Sesemble by: 1st > part of name, separated by (Hyphen) × > Assemble each paired read separately Use 100 \$ % of data. Suitable for genome size between 0 KB and 0 KB. Method Assembler: Geneious ? ? Not sure which assembler to use? Let us help! Sensitivity: Medium Sensitivity / Fast ? ? Mermory Required: 84 MB of 13 GB Note: Paired reads can be sequence. Net: Paired reads can be sequence. Save assembly neport Save assembly report Save assembly report Save assembly report Save assembly report Save in sub-folder Ø Save consensus sequences Save consensus sequences Options	Edit Align/Assemble -> De Novo Assemble				
Deptions to expose to user when workflow is run Expose all options Expose all options Expose some options Optionally label exposed options as: Expose: None VII Operation Options (those not exposed to workflow user and default values for options that are exposed) Data I Operation Options (those not exposed to workflow user and default values for options that are exposed) Data I Ossemble by: I st part of name, separated by -(Hyphen) Assemble each sequence list separately Use into % with Atternative Label: I Assemble each sequence list separately Use into % with assembler to use? I Assemble each sequence list separately I Assemble each sequence list separately I Assemble each sequence list separately I Sensitivity: Method Assembler: Ceneous ? Not sure which assembler to use? I Sensitivity: Memory Required: 84 MB of 13 GB Note: Paired reads can be set up or changed using Sequence > Set Paired Reads Trim Before Assembly I Save existing trim regions I Use existing trim regions from sequences I Save construct of unused reads Save list of unused reads Save consensus sequences I Save consensus seque	Assemble reads (eg. Sanger or NGS) without using a reference				
Expose no options Expose all options Optionally label exposed options as: Access exposed options via button ? Expose: None With Alternative Label: (Optional) + - VII Operation Options (those not exposed to workflow user and default values for options that are exposed) Data I Dissolve contigs and re-assemble I Dissolve contigs of data. Suitable for genome size between 0 KB and 0 KB. Method Assembler: Geneious ? ? Memory Required: 84 MB of 13 GB Note: Paired reads can be set up or changed using Sequence > Set Paired Reads Trim Before Assembly I Dissolve contigs from sequences I Save assembly report Save assembly report Save in sub-folder I Dis out rim (discard trim annotations) I Dissolve folder I Dissolve contigs (I Maximum 1,000 I) Save consensus sequences Options I Dissolve con	Options to expose to user when workflow is run				
Expose all options Expose some options Optionally label exposed options as: Expose: None Vith Alternative Label: (Optional) Expose: None Vith Alternative Label: (Optional) Home With Alternative Label: Upperation Options (those not exposed to workflow user and default values for options that are exposed) Data Image: Data	 Expose no options 				
○ Expose some options Optionally label exposed options as: Expose: None ✓ With Alternative Label: (Optional) Expose: None ✓ With Alternative Label: (Optional) H - All Operation Options (those not exposed to workflow user and default values for options that are exposed) Data ✓ Dissolve contigs and re-assemble ✓ Seemble by: 1st ● assemble each sequence list separately ● Assemble each sequence list separately ● use 100 ○ % of data. Suitable for genome size between 0 KB and 0 KB. Method Assembler: Geneious ○ ? ? Method Assembler: Sensitivity: Medium Sensitivity / Fast ○ ? Memory Required: 84 MB of 13 GB Note: Paired reads can be set up or changed using Sequence > Set Paired Reads Trim Before Assembly Results Assembly Name {Reads Name} Assembly ✓ Use existing trim regions Save assembly report Save assembly report Save in sub-folder Ø ave constigs Y Maximum 1,000 ○) Ø ave constigs Options <td> Expose all options </td> <td></td>	 Expose all options 				
Optionally label exposed options as: Expose: None All Operation Options (those not exposed to workflow user and default values for options that are exposed) Data Isolate Ist part of name, separated by - (Hyphen) Isolate	Expose some options				
Expose: With Alternative Label: (Optional) + - All Operation Options (those not exposed to workflow user and default values for options that are exposed) Data I a Dissolve contigs and re-assemble I semble by: I st v part of name, separated by - (Hyphen) × v Assemble each sequence list separately Assemble each paired read separately Use 100 % of data. Suitable for genome size between 0 KB and 0 KB. Method Assembler: Geneious ?? Not sure which assembler to use? Let us help! Sensitivity: Medium Sensitivity / Fast ??? Memory Required: 84 MB of 13 GB Note: Paired reads can be set up or changed using Sequence > Set Paired Reads Trim Before Assembly I use existing trim regions O use existing trim regions from sequences Save assembly report Save assembly report Save in sub-folder Save contigs (2 Maximum 1,000 ?) Save consensus sequences Options	Optionally label exposed options as:	Access exposed options via button ?			
All Operation Options (those not exposed to workflow user and default values for options that are exposed) Data	Expose: None	V With Alternative Label: (Optional)			
♥ Dissolve contigs and re-assemble ● seemble by: 1st > part of name, separated by -(Hyphen) × > ● Assemble each sequence list separately ● Assemble each paired read separately ● Use 100 ◇ % of data. Suitable for genome size between 0 KB and 0 KB. Method Method Assembler: Genelous > ? ● Not sure which assembler to use? Let us help! Sensitivity: Medium Sensitivity / Fast > ? Memory Required: 84 MB of 13 GB Note: Paired reads can be set up or changed using Sequence > Set Paired Reads Trim Before Assembly ● Use existing trim regions ● Use existing trim regions from sequences ● Remove existing trim regions from sequences ● Save in sub-folder ● Do not trim (discard trim annotations) ● Do not trim (discard trim annotations)	All Operation Options (those not exposed to wor Data	kflow user and default values for options that are exposed)			
Issemble by: 1st v part of name, separated by - (Hyphen) × v Assemble each sequence list separately Assemble each paired read separately Use 100 % of data. Suitable for genome size between 0 KB and 0 KB. Method Assembler: Geneious v ? Not sure which assembler to use? Let us help! Sensitivity: Medium Sensitivity / Fast ? Memory Required: 84 MB of 13 GB Note: Paired reads can be set up or changed using Sequence > Set Paired Reads Trim Before Assembly Results Assembly Name {Reads Name} Assembly Vise existing trim regions Do not trim (discard trim annotations) On ont trim (discard trim annotations) On the pair of name, separated by - (Hyphen) × v Output:	✓ Dissolv	e contigs and re-assemble			
Assemble each sequence list separately Assemble each paired read separately Use 100 ♀ % of data. Suitable for genome size between 0 KB and 0 KB. Method Assembler: Geneious ♥ ? Not sure which assembler to use? Let us help! Sensitivity: Medium Sensitivity / Fast ♥ ? Memory Required: 84 MB of 13 GB Note: Paired reads can be set up or changed using Sequence > Set Paired Reads Trim Before Assembly Results Assembly Name {Reads Name} Assembly Subse existing trim regions Save assembly report Save list of unused reads Save in sub-folder Save contigs { Maximum 1,000 ♀ } Save contigs { Maximum 1,000 ♀ } Save contigs { Maximum 1,000 ♀ }	ssemble by: 1st	\sim part of name, separated by $ $ - (Hyphen) $ imes imes imes$			
Use 100 ◇ % of data. Suitable for genome size between 0 KB and 0 KB. Method Assembler: Geneious ◇ ? Image: Second Sec	Assemble each seq	uence list separately Assemble each paired read separately			
Method Assembler: Geneious V ? Not sure which assembler to use? Let us help! Sensitivity: Medium Sensitivity / Fast V ? Memory Required: 84 MB of 13 GB Note: Paired reads can be set up or changed using Sequence > Set Paired Reads Trim Before Assembly Results Results Assembly Name {Reads Name} Assembly Save assembly report Save list of unused reads Save in sub-folder Save in sub-folder Save consensus sequences Options Save consensus sequences Options	Use 100 🔇	% of data. Suitable for genome size between 0 KB and 0 KB.			
Sensitivity: Medium Sensitivity / Fast Memory Required: 84 MB of 13 GB Note: Paired reads can be set up or changed using Sequence > Set Paired Reads Trim Before Assembly Results Note: Paired reads Subse existing trim regions Remove existing trim regions from sequences Re-trim sequences Options Do not trim (discard trim annotations) Sensitivity: Memory Required: 84 MB of 13 GB Note: Paired reads can be set up or changed using Sequence > Set Paired Reads Assembly Name {Reads Name} Assembly Save assembly report Save is to of unused reads Save in sub-folder Save contigs (Maximum 1,000 \$) Save consensus sequences Options		Not sure which assembler to use? Let us help!			
Memory Required: 84 MB of 13 GB <i>Note: Paired reads can be set up or changed using Sequence > Set Paired Reads</i> Trim Before Assembly Results O Use existing trim regions Assembly Name {Reads Name} Assembly Vuse existing trim regions from sequences Save assembly report Save list of unused reads Save in sub-folder Save contigs (Save contigs (Maximum 1,000 \$) Save consensus sequences	Sens	itivity: Medium Sensitivity / Fast 🛛 🗸 ?			
Note: Paired reads can be set up or changed using Sequence > Set Paired Reads Trim Before Assembly Results Assembly Name {Reads Name} Assembly Save assembly report Save list of unused reads Save list of unused reads Save in sub-folder Save contigs (Save contigs (Maximum 1,000 \$) Save consensus sequences Options		Memory Required: 84 MB of 13 GB			
Trim Before Assembly Results Assembly Name {Reads Name} Assembly Assembly Name {Reads Name} Assembly Save assembly report Save list of unused reads Save in sub-folder Save contigs (✓ Maximum 1,000 ◊) Save consensus sequences Options	Note: Paired read	is can be set up or changed using Sequence > Set Paired Reads			
 Assembly Name {Reads Name} Assembly Assembly Name {Reads Name} Assembly Save assembly report Save list of unused reads Save in sub-folder Save contigs (Maximum 1,000 \$) Save consensus sequences Options 	Trim Before Assembly	Results			
O Use existing trim regions ✓ Save assembly report ○ Remove existing trim regions from sequences ○ Save list of unused reads ○ Re-trim sequences ○ Options ○ Do not trim (discard trim annotations) ✓ Save contigs (✓ Maximum 1,000 ◇) ○ Save consensus sequences ○ Options		Assembly Name {Reads Name} Assembly			
 Remove existing trim regions from sequences Re-trim sequences Options Do not trim (discard trim annotations) Save contigs (✓ Maximum 1,000 ◊) Save consensus sequences Options 	Use existing trim regions	Save assembly report			
Re-trim sequences Options Do not trim (discard trim annotations) Save consensus sequences Options	Remove existing trim regions from sequence	es Save list of unused reads			
 ○ Do not trim (discard trim annotations) ○ Save contigs (✓ Maximum 1,000 ◇) ○ Save consensus sequences Options 	Re-trim sequences Options	Save in sub-folder			
Save consensus sequences Options	 Do not trim (discard trim annotations) 	Save contigs (V Maximum 1,000 🗘)			
		Save consensus sequences Options			

この設定では、ファイル名のダッシュ(ハイフン)の前に同じ文字列を持つクロマトグラムを 1 つのコン ティグにアセンブルしますので、今回の例では、トリミングされた領域を除く各サンプルのフォワードリ ードとリバースリードをそれぞれアセンブルしてコンティグをアウトプットします。

もし、この段階で複数のアウトプット(例えばコンティグとコンセンサスシークエンスなど)を設定した場合には、次の解析ステップには、フィルタリングオプション(Add step → Filter documents)な どを使用して、解析に使用するアウトプットを指定する必要がありますのでご注意ください。

同様に、

- Generate Consensus Sequence
 Alignment → MUSCLE Alignment
- の2つの操作を追加します。この例ではオプションはデフォルトのままです。

• • •	Create Workflow		
Workflow Name:	Sanger analysis		
Author:	Geneious Prime		
Description:	Sanger sequences trimming, de novo assembly, multiple alignment and tree building		
Icon:	👗 Chromatogram 🗸 Choose Custom Icon		
Ô	Share (read-only) with other Shared Database users		
ß	Share (read-only) with other Geneious Cloud users		
+ Add Step $-$	Delete Step $~ \slash $ View/Edit Options $~ \hfill hove Up ~ \lash $ Move Down $@$ Help		
▶ Trim Ends Options: Error probability=0.01; Trim regions ▶ Align/Assemble -> De Novo Assemble Options: 'No Documents'; Trim regions ▶ Generate Consensus Sequence ▶ Alignment -> MUSCLE Alignment			
	Cancel OK		

最後のステップとして、ワークフローに以下を追加します。

TreeBuilding → Geneious Tree Builder (nucleotide)

ワークフローの動作時に、系統樹の構築方法(Neighbor-Joining または UPGMA)を変更するオ プションを変更できるようにするため、Expose some options を選択し、Tree building Method オプションを設定します。

次回はワークフローを完成させ、動作を確認する流れについてご紹介します。

Geneious 製品概要・フリートライアルリクエストについては<u>こちら</u> 『Geneious Prime で猫も杓子もシークエンス解析』 過去の記事は<u>こちらでチェック!</u>