

DDBJ Read Annotation Pipeline の紹介と実習

国立遺伝学研究所 大量遺伝情報研究室

長崎 英樹

長崎は遺伝研 大量遺伝情報研究室の所属です。



私たちは 塩基配列登録を支援するシステムづくり 登録データの活用するシステムづくり 高速シーケンス配列の情報解析 を行なっています。

1

高速シーケンサー配列の登場で短期間、低コストで大量 の塩基配列データを出力されるようになった。



illumina社 HiSeq2000



X	
	-
Roche社	GS FLX



Life Technologies社 SOLiD

データ保管場所の確保

計算機不足

解析のための人員不足

といった問題がでてきた。

DDBJの高速シーケンサー配列の諸問題への対応 DDBJ DRA INSDC **DDBJ** Sequence データ保管場所 NCBI FNA/FB Read Archive (DRA) の確保 SRA FRA International Nucleotide Sequence >中村、児玉の発表をご参考 Database Collaboration (INSDC) ください。

計算機不足
NIG(遺伝研)スパコンシステム
O アカウント登録で無償利用
X コマンドラインによる操作
X データ規模や使用メモリ量等で計算機ノードを選択などコツがいる。

解析のための 人員不足



DDBJ パイプラインの特徴

・遺伝研の計算機で分散処理を実行、高速シーケン スデータを解析するクラウド型パイプライン

・オンラインで無償で利用可。

・基礎解析部 (マッピング、*de novo* アセンブル)と 高次解析部 (構造・機能のアノテーション)で構成





2			2230	panese
DDBJ Read Annotation Pipeline is a cloud-con data.	nputing based analytical plat	tform for next-g	eneration sequencing	
LOGIN		New a	ccount Login as "gu	est".
Pipeline Flow UST Pipeline Flow COEI Read Active CoEI Rea	User ID: Password:	Login		
Areadon mage pesitiene WES/CON Basic Analysis Mapping de novo atsenday	Create a new act	count. ock current jobs.		
High-level Analysis	Data Transfer	File Format	Menu Item	
Bineline	1 DRA(see HP)	FASTQ	DRA Start	
Pipeime	2 FTP Upload	FASTQ/FASTA	FTP Upload	
Operation Display We changed the accessible letter number of login passwords of DDBJ Pipeline from 9 to 16. 17 days ago - reply - retweet - favorite 11/ days ago - reply - retweet - favorite 19 days ago - reply - retweet - favorite 19/ bags ago - reply - retweet - favorite 19 days ago - reply - retweet - favorite DDBJ pipeline is now available : system maintenance was finished. 27 days ago - reply - retweet - favorite 11/ hoticell11 DDBJ pipeline services will no be available due to system 10 in the conversation	Manual & tutoi Japanese, ma English man English man English man English man English man English man State of the second	rial insual ial insual. (JP) v Tutorial video of DDBJ Read istics of monthly res of pipelines: stics of monthly rration of "DRA	1 (JP) - Reference 2 (JP) - De novo Annotation Pipeline olpeline's web acces disc usage, pipeline's job, *	
Ewittery Join the conversation	page.			

DDBJパイプライン 基礎解析部

http://p.ddbj.nig.ac.jp

・11種類のマッピング・アセンブルソフト対応

マッピング

BLAT	高速シーケンサー登場以前からあるアライメントツール。 発現データはイントロンを想定したギャップを考慮。
MAQ	高速シーケンサー登場初期にショートリードに対応。 リード長が長くなるに従い開発はBWAに引き継がれる。
BWA	MAQより速く、Titaniumのリードもオプションで対応。
SOAP	メモリ消費量少なく、より高速、精度はBWAより弱冠落ちる。
Bowtie/ Bowtie2	ギャップは考慮しないが処理は速い。BWA、SOAP2、Bowtieは Burrows-Wheeler変換というアルゴリズムでゲノムDNAにたいし てインデクスを作成、高速でマッピングする。Bowtie2は50bp以 上に最適化。
TopHat	RNA-Seqのリードを内部でBowtieを利用してマッピング、スプ ライスジャンクションを特定する。

アセンブル

SOAPdenovo	ヒト、パンダ等大型ゲノムのアセンブリで使用された。比較的高 速。
Abyss	初期に並列処理に対応したアセンブラ。
Velvet	高速シーケンサー登場初期に開発された。メモリ消費多め。
Trinity	RNA-Seq配列のアセンブラ。 上記3つともにde bruijn graphと いうアルゴリズムを使用。

0	Refere	ance	Gen	ome Ma	apping	6									(BACK) (NEXT
				Input d	lata		Evalua	tion		Anal	ysis	Out	put fo	rmat	
	Tool	Help	Versi	on Base space	Color space	Paired	Depth	Coverage	Error	SNP	Indel	.gff	.bed	SAM	Comment
9	BLAT		34	v					v						Single-end analysis only
3	Mag :?	4	0.7.1	¥		¥			v	v	v	v	V)	¥.	
3	DWR.(7	4	0.5.9						v					v	
3	SOAP.	•	2.21	v		*			v	v	¥			×	
3	Bowle,	*	0.12.3		w.,	4			v.	v.				¥.	
a	TopHat		1.0.11			*			¥					*	
-	Total la	nit = 2	12 Gbp Help	Version	Base space	Color space	Paire	d- MSS(W	GS)	Com	ment				
	SOAPden C	<u>010</u>	9	1.05											
	ABySS./	1	0	1.3.2			¥					Mar	emun	n K-me	r value is 64.
	Velvetra		•	1.2.03			v	v		We slengt	ievera h of th	record ose n	mmer eads	d whe	n performing Velvet, total 22G bp Maximum K-mer s 64.
1	Trinity. ¹⁷		•	r2012- 06-08			*					RNA	A-Seq	Den	wo Assembly
	Mappi The con	ng C	il be a	gs by d	e novo	Asse	mble	to Refe	rence	s Se	quer	ices	s.,		
1															

DDBJパイプライン 基礎解析部

- ・11種類のマッピング・アセンブルソフト対応
- ・公開配列データの活用が容易 公開データと比較、レファレンスとしての活用

	Refere	ence	Ge	nome Mi	pping										(BROK) (HEXT
				Input d	ata		Evaluat	ion		Anal	ysis	Out	put fo	rmat	
	Tool	Help	Ver	ion Base space	Color space	Paired end	Depth (Coverage	Error	SNP	Indel	.gtt	.bed	SAM	Comment
۲	BLAT	4	34	×	200				¥						Single-end analysis only
۵	Mag	•	0.7.1	×		4			*	¥		¥	¥	*	
۵	twa 🕫	4	0.5.9	¥ 🖌		*			*					*	
0	SOAP D		2.21	~		×			v	v	×			*	
0	Bowtie.	•	0.12	7 🖌		~			×	v				v	
0	TopHat	4	1.0.1			v			v					v	
	Tool		Help	Version	Base space	Color space	Paired end	MSS(W	GS)	Com	nent				
٥	SOAPden 17	21/2		1.05			×								
3	ABySS:?		•	1.3.2			¥					Ma	dmun	n K-me	ar value is 64.
•	Velvet.c?		9	1.2.03			*	~		We slengt	evere h of th	neco ose n	mmen eads i	nd whe is up to value i	en performing Velvet, total o 22G bp Maximum K-mer s 64.
	Trinity.c?	1	\$	r2012- 06-08			۲					RN	A-Seq	Deno	ovo Assembly
	Mappi The con	ng (Cont III be	igs by d aligned to n	e novo	Asse	mble t	o Refe	rence	e Se	quer	ces			
	-				11										

DDBJパイプライン 基礎解析部

http://p.ddbj.nig.ac.jp

- ・11種類のマッピング・アセンブルソフト対応
- ・公開配列データの活用が容易

公開データと比較、レファレンスとしての活用

・ジョブステータスで実行状態を確認可能

NIGスパコンで実行 マッピング Intel Xeon 2.60GHz 16 core,64GB RAM * 352 nodes

アセンブル

Intel Xeon 2.40GHz 80 cores, 2TB RAM * 2 nodes Intel Xeon 2.66GHz 768 cores, 10TB RAM

ストレージ 2PB storage

解析終了をメールで通知

・SAMtools/FASTAによる共通フォーマット での出力

Sel	ect Que	ry film	Belect Tools	let C	werythet	+ Set Ce	nomeBet)++(Set Ma	p Options)-+ (Cor	demation)	1
Hur	anng a	22.4										
SI	latu	s - Mapp	ing									
				-	oping gues	y state >>	Assembly qu	ery stat	•••)(PreProcess	query state >	•
Ord	er	-	173									
orti	by: lat		Descending 10	only	the login u	ser (foat	000					
De	iete)						2 3 4 5 6	7 8	0 10 11 21 22 2	12 13 3 24 25	14 15 16 26 27 28	17
	ю	UserID	Submission accession	P/5	Status	Tool	Read #	Read	Genome size	Download	Start time End time	Elapsed time
	3233	ekaminuma	SRA009756 Mo17	5	complete	trea	49,285,631	1	4,185 M	(10)	2011-12-10 13:24:27	00:28:54
			Same in the								2011-12-10 13:53:22	
	3224	nagasakicool	DRA000369 kyctou_pmb-00	P	complete	DWB	122,403,348	-	228 M		2011-12-08 16:07:20	63:54:02
2										Fie	2011-12-11 08:01:23	
	3220	-	ERA000212 Glad-L2 Mount	P	complete	bes	2,873,320	40	83 M		2011-12-07	00.29.37
9										fit	2011-12-07	
	3217	user-demo	SRA000284 GSM276809.1	8	complete	bwa	28,723,026	- 77	2,708 M	_	2011-12-07 13:21:20	00.21.13
2										(File)	2011-12-07	
	3215	user-demo	5RA030871 newly_synthesic	5	complete	TopHat	4,671		254 M	Sec. 1	2011-12-07	00:17:36
2										(File)	2011-12-07	
	3254	user-demo	DRA000307 tomohiro-0005	p	complete	bwa.	17,359,151	-	308 M		2011-12-07 10:13:34	08-33-30
3			- 3							File	2011-12-07 18:47:04	-
	3212	tmochidu	DRA000158	P	complete	bws	50,531,840	110	390 M	1	2011-12-06	19:39:07



DDBJパイプライン 高次解析部 http://p-galaxy.ddbj.nig.ac.jp

NGSデータのマッピング結果の解析

SNPのゲノム上の分布の表示

<complex-block>

RNA-SeqのCufflinks実行(発現量の正規化) gtf->wigフォーマット変換 UCSC genome browser siteでの可視化



⁽http://genome.ucsc.edu/cgi-bin/hgGateway)

ChIP-Seq MACSによるDNA結合タンパク質の結合部位候補の同定

DDBJパイプライン 高次解析部 http://p-galaxy.ddbj.nig.ac.jp

RNA-Seqのde novo アセンブル結果の解析



DDBJパイプラインで実行するTrinityについて

Inchworm: k-mer(k=25)でざっくりアセンブルしてコン ティグをつくる。

Chrysalis:

スプライスバリアントやパラログ由来のコン ティグを含めてクラスター化 コンティグの共通部分を基にどういう経路を とってつながっていくか? >グラフを作成

Butterfly:

グラフを精査していってスプライスバリアント やパラログも再構成する。

Trinityについては Nat Biotechnol. 2011 May 15;29(7):644-52. グラフアルゴリズムについては http://d.hatena.ne.jp/hoxo_m/20100930/p1 等ご参考ください。



Figure 1 Overview of Trinity. (a) Inchworm assembles the read data set (short black lines, top) by greedily searching for paths in a *k*-mer graph (middle), resulting in a collection of linear contigs (color lines, bottom), with each *k*-mer present only once in the contigs. (b) Chrysalis pools contigs (colored lines) if they share at least one k – 1-mer and if reads span the junction between contigs, and then it builds individual de Bruijn graphs from each pool. (c) Butterfly takes each de Bruijn graph from Chrysalis (top), and trims spurious edges and compacts linear paths (middle). It then reconciles the graph with reads (dashed colored arrows, bottom) and pairs (not shown), and outputs one linear sequer \overline{gre} for each splice form and/or paralogous transcript represented in the graph (bottom, colored sequences).

今回はミドリフグのRNA-Seqデータを使用します



Tetraodon nigroviridis

最大で全長17 cm。 観賞魚としてポピュラーであり、2-3 cm程度の幼魚 が多くの熱帯魚店等で売られている。

SRR042533 (エントリー: SRA012701) 36bpの7,468,448リード シングルエンド





大量遺伝情報研究室の方々 富士ソフト株式会社 森崎さん

DDBJの方々

本研究は、文部科学省科学研究費新学術領域研究『生命科学系3分野支援活動』 「ゲノム支援」および科学研究費基盤(C)の支援を受けております。

大量研ではDDBJパイプラインをカンキツ類、野生イネ、ミニトマト、ゼニゴケ等の変異解析、パラゴムの木のアセンブルに使用しております。

DDBJ Read Annotation Pipeline: a cloud computing-based pipeline for high-throughput analysis of next-generation sequencing data. DNA Res. 2013 Aug;20(4):383-90.

実習内容

DDBJ パイプラインを用いた denovo RNAseq アセンブリ

DRA (DDBJ Sequence Read Archive)からの配列データのインポート
DDBJパイプライン基礎部での Preprocessing ジョブ実行
DDBJパイプライン基礎部での Trinity ジョブ実行
DDBJパイプライン高次解析部(Galaxy)でのジョブ実行

参考資料

DDBJパイプライン(基礎部)へのアカウント作成 DDBJパイプライン(基礎部)のFTPによるデータ転送

DRAからの配列データインポート

今回使用する高速シーケンサー配列の確認

DRAで検索すると早い

今回は実習用サンプルとしてミドリフグの高速シーケン サーで出力された RNAseq 配列を用いる。

DRAのwebサイトから「検索」をクリック

「Organism:」に「Tetraodon nigroviridis」

今回はアクセッション番号「 SRA012701 」のデータ

に必要なので、アクセッションをメモしておく。

DRASearchのwebサイトが表示

と入力し、「Search」をクリック。

DRA: http://trace.ddbj.nig.ac.jp/dra English Login & Submit Sitemap Contact Sequence Read Archive Google"カスタム機関 Q, Home Submission - Search Download - Pipeline

2013年04月25日: ウェブサイトをリニューアルしました mon DDBJ Sequence Read Archive (DRA) は Roche 454 GS System®, Illumina Genome Analyzer®, Applied Biosystems SOLD® System などの次世代シー クエンサからの出力テータのためのテータベースです。 DRA は International Nucleotide Sequence Database Collaboration (INSDC) のメンバーであり, NCBI Sequence Read Archive (SRA) と EBI Sequence Read Archive (ERA) との国際協力のもと、運営されています。 従来のキャピラリ式シークエンサから の出力データは DDBJ Trace Archive にご登録ください。 0 **영양** ダウンロード 検索 新型シークエンサからの生 データやアライメントデー タを登録する メダデータや配利データを ファイルとして ftp ダウン ロードする DRASearch Send Feedback * Search Home * DRA Home Tetraodon nigroviridis dyType : - 1 word : how 20 a records Sort by Study Search Results (4 studies) << < 1 / 1 Page > >> BASES SUBMITTED CENTER NAME # STUDY TUDY_TITU STUDY_TYPE をサンプルに用いる。Pipelineからインポートするの SUBMISSION 2010-07-21 GEO SRA012701 5.1G 2 SRP002418 SRA012201 GSE19824: Genome-wide evolutionary analysis of eukaryotic DNA methylation (ChIP-Sea) 1.3G 2010-07-21 GEO Epige

DDBJパイプラインにログイン





http://p.ddbj.nig.ac.jp/

DDBJ, pipeline で検索すると早い

デモ用アカウントは 講習内でお伝えします

http://www.ddbj.nig.ac.jp/

DRAから配列データをインポート

	DDDBJ DXA Date Reek of Jopen	Select Que	ry Files → S	Select Tools Set QuerySet	Set GenomeSet	Set Map Options	Confirmat	tion →
DDBJパイプラインログインする。	Iogin ID [guest]	Select	ting Que	ry Files				
「Import public DRA」をクリック	ANALYSIS Data setup DRA Start	FTP uplo	ad Privat	e DRA entry Import publ	選択 ic DRA Prep	rocessing HTTP	upload	NEXT
	FTP upload HTTP upload DRA Import Preprocessing Start	Metadat	a of the DRA er	ntry.		Select a me	tadata : DRA0	00001
	sten-1	TYPE	ACCESSION	ALIAS	FI		DI	VIEW
	Preprocessing	Puteriosian	DBA000004		0.0	A000001 submission uml	(Deveload)	(Marr)
	Mapping / de novo Assembly step-2	Sample	DRS000001	Bacillus subtilis subsp. natto BEST plasmid pBEST195L	195 without DR	A000001.sample.xml	DownLoad	View
	Workflow	Study	DRP000001	Natto BEST195	DR	A000001.study.xml	(DownLoad)	(View)
	Genome (SNP/Short Indel) BNA-sec (Tag count)	Experiment	DRX000001	NATTO_BEST195_SEP08	DR	A000001.experiment.xml	(DownLoad)	View
	ChIP-seq	Run	DRR000001	2008-09-12.BEST195-Lane7	DR	A000001.run.xml	(DownLoad)	(View)
「Input DRA/ERA/SRA Accession Number」 「SRA012701」と入力 「Add my DRA entry」をクリック	رت Selecting	g Query	Files	Import public DRA	Preprocess	sing HTTP uple	oad	NEXT
	Import public	FASTO files f	rom DRA dat	tabase.				-
	Here is do the sec Please input DRA	tion of automa VERA/SRA ac at DRA/ERA/S SRA012 sion Number of iearch	atic download ccession nun RA Accession 701	of public DRA/ERA/SRA entrie nber. Then the pipeline system n Number Add my DRA entry	s. import metadata リック	and FASTQ files from I	DRA database	B.

DRAから配列データをインポート

「Confirmation」のダイアログが現れる。

<code>'Send a mail when completed importing]</code> \mathcal{OF}_{\pm} を確認。チ

Send a mail when completed importing 07 177		This operation may take several r	minutes to several hours.	
を確認。チェックしておくとimport終了時にメールが届く。	terres a	Option		
	E F	Send a mail when completed	importing.	
		Show a accessions list.		
$\int OK_{\perp} \langle F \rangle D \langle V \rangle D \langle V \rangle$	Your re			
	To select y			
	When the s When the s		クロック	7
	queuec		////	ry it,
	Status		ОК	Cancel
	en done			1.
	Here is do the section of auto Please Input DRA/ERA/SRA	matic download of public DRA/ERA/SRA accession number. Then the pipeline s	A entries. system import metadata and FAST	Q files from DRA database.
webノノクリをクロードして「力の八子クストを確認。		Add my DRA entry		
実行中のDRAのアクセッションが「queued」から「done」				
になったら完了。	Accession Numbe	r can find here.		
	MON SHRIDD			
	Your request. (Here Is di	splay only. can not select.)		
	Your request. (Here is di To select your downloaded e When the status makes "doine When the status makes "daile	splay only, can not select.) ntries. See Private DRA entry tab. " your requested entry is added in "Priv. " or "preparing", please retry it.	ate DRA entry" tabs.	
	Your request. (Here is di To select your downloaded e When the status makes "done When the status makes "done When the status makes "done URA unchecked ::	splay only. can not select.) ntries. See Private DRA entry tab. ", your requested entry is added in "Priv " of "preparing", please retry it. ring download, dome : file is ready wenikad is do, but mdd was notcheck.	ate DRA entry" table. y, falled :please retry it,	preparing : file is not yet in
ブラウザリロードで確認	You request. (Here is di Your request. (Here is di To select your downloaded e When the status makes Table When the status makes Table gueued : waiting or du DRA unchecked : c	splay only. Can not select.) ntries. See Privata DRA entry tab. ", your requested entry is added in "Priva or preparing", please netry it. ining download, donefile is ready wenned as in ab, but md5 was not check. Submission	ale DRA entry" tabs. y, falled : please retry it, Request date	preparing : file is not yet in

11

FTP upload Private DRA entry Import public DRA Preprocessing

ease input DRA/ERA/SRA accession number. Then the pipeline system import metadata and FA:

re is do the section of automatic download of public DRA/ERA/SRA entries

Import public FASTQ files from DRA database.

Click a OK button to start import.

Preprocessing リードのクオリティ値によるフィルタリング

Preprocessing 実行するクエリファイルを選択

Trinity 実行の前に、インポートしたデータの前処理として、QV によるフィルタリングを行う



ウィンドウ下部にメタデータおよびファイル一覧が表示されるので、 この中から、Tetraodon_nigroviridis_RNA-Seq に該当する Experimental ACCESION20122 のものをチェック。

sin	gle	paired	all	dear							
	No.	Experim	sion	Sample ACCESSION	Run ACCESSION	STRAIN	Run_date	Read #	Read length	Instrument model	Layou
0	1	SRX0201	112	SRS070561	SRR042523	strain n/a				ILLUMINA	single
Ó	2	SRX0201	113	SRS070562	SRR042524	strain n/a				ILLUMINA	single
0	3	SRX0201	114	SRS070563	SRR042525	strain n/a				ILLUMINA	single
0	4	SRX0201	115	SRS070564	SRR042526	strain Okayama 7				ILLUMINA	single
	5	SRX0201	116	SRS070565	SRR042527	strain ATCC #64925				ILLUMINA	single
0	6	SRX0201	117	SRS070566	SRR042528	strain n/a				ILLUMINA	single
0	7	SRX0201	18	SRS070567	SRR042529	strain japonica cultivar Nipponbare				ILLUMINA	single
	8	SRX0201	119	SRS070568	SRR042530	strain NRRL 1555				ILLUMINA	single
Ð	9	SRX0201	120	SRS070569	SRR042531	strain ATCC #11538				ILLUMINA	single
	10	SRX0201	121	SRS070570	SRR042532	strain n/a				ILLUMINA	single
3	11	SRX0201	122	SRS070571	SRR042533	strain n/a				ILLUMINA	single
0	12	SRX0201	123	SRS070572	SRR042534	strain #UAMH 1704				ILLUMINA	single
0	13	SRX0201	124	SRS070573	SRR042535	strain UTEX #LB 1885				ILLUMINA	single
0	14	SRX0201	125	SRS070574	SRR042536	strain ATCC #50258				IL12MINA	single

Preprocessing 実行条件の指定

Trinity 実行の前に、インポートしたデータの前処理として、QV によるフィルタリングを行う

	Read length	Quality Score	Read Layout
RR042533 ->	bp		single
teps of prep	processing	workflow	
Step1: Set th	e type of the q	uality value.	
0		ク	オリティ値の選抜
 Phred 	+33 🔘 Phre	$d+64 \rightarrow \overline{\tau}$	ータはすべて P
If you don'	t know it, pleas	e see <u>'2.2 Encoc</u>	ling' of this site
Step2: BASE	TRIMMING with	th low quality fro	om 5'end and 3'e
Bases with bases of the If read len	n low quality (C ne trimmed rea gth after base t	QV <= THRESHO d indicate high q rimming is too sh	DLD) are trimmed uality (QV > THF nort (length <= 24
length will	be 25bp.	リート	、の両端から Q
 QV THRE 	SHOLD :	19 → トリ (ペ)	ム後の長さが 25 Pの場合は、ペご
Change DEAD	REMOVING to	o discard trimm	ed reads includi
Steps: READ		normaniana (h-	Low quality base
Trimmed r THRESHO	eads with high DLD) are disca	rded.	
Trimmed r THRESHO	eads with high DLD) are disca	rded.	トリム後の
Trimmed r THRESHO	eads with high DLD) are disca SHOLD :	rded.	トリム後の → 以上含まれ

Preprocessing 実行および実行状況の確認

Trinity 実行の前に、インポートしたデータの前処理として、QV によるフィルタリングを行う

T

			Email	notification							
メールを入力して「Run」ボタンを押す。			Send email	notification when the job is o	complet	ed or aborted wi	th error.				
			Confir	mation of entries							
			Query s	ots SRR042533 - GSM497271_	_1						
										B	ACK RUN
ステータス画面でジョブの実行状況の確認。	S	tatus	- Prep	processing							
							Mappir	ig Job	de novo	Assembly Job	Preprocessing Job
						1					
	Sort	er bv: ID	•	Descending 1	Show	v Only Your (Own Job	Relo	ad		
	0011	oy				,		Citato			
Preprocessing でフィルタリングをした	Del	ete *									page 1 ÷ NEXT >
クエリファイルを利用してdenovo Assemblly		ID	UserID	Files	P/S	Status	Read #	Read length	Detail	Start time End time	Elapsed time
/ mapping を行う場合、ジョブエロが必要になる		5509	toshu01	SRA012701 GSM497271_1	S	running		-	View	2013-04-30 17:42:30	
1									view		
りて、見んておくこと。		5455		FY23KIH080_pl	P	complete				2013-04-25 11:55:45	01:20:10
				(Th)						2013-04-25 13:15:56	
	1265	5452		 FY22KIH033_pl	P	complete				2013-04-25 11:16:44	01:27:43

「View」ボタンで詳細を確認。

2013-04-25

Trinity 実行の前に、インポートしたデータの前処理として、QV によるフィルタリングを行う



「BACK」ボタンでジョブ履歴画面に戻る

denovo Assembly Trinity の実行

Phred Quality Score

Trinityの実行 クエリファイルの選択

クエリとなるFASTQ/FASTA配列を選択する方法としてDDBJパイプラインでは、下記の4通りの方法がある。

• FTPクライアントソフトでアップロードした配列を使用

^rFTP upload」

- webブラウザでアップロードした配列を利用 「HTTP upload」
- DRAからインポートした配列を使用する 「Private DRA entry」
- Preprocessing で処理した配列を使用 「Preprocessing」

FTP upload	Private DRA entry	Import public DRA Preproce	HTTP uploa	ad		
From Prepro	cessing output files.					
Filename			Layout	File size		
4927_DRR0	000719_1.unmapped.fastq_4	paired	65.8 ME			
4932_DRR0	00719_1.unmapped.fastq_4	746.bz2 (more 1 files)	paired 6			
4971_DRR0	00719_1.unmapped.fastq_4	785.bz2 (more 1 files)	paired	65.8 MB		
1 5509_SRR0	42533_e.fastq.bz2		single	239.7 ME		

選択

次へ

今回は Preprocessing で処理したクエリを使用する。 画面左のメニューから、「Preprocessing Start」を選択。

Preprocessing で処理されたファイルは、

「(PreprocesingのジョブID)_もとのファイル名_e.fastq.bz2」という形式のファイル名になっているので、 先ほど確認しておいたジョブIDで始まるものを選択。

最下部の「NEXT」をクリック。

Trinityの実行 ツールの選択

「denovo Assembly」 → 「Trinity」の順に選択

•	de novo Asa Total limit =	semb 22 Gbp	ly					
	Tool	Help	Version	Base space	Color space	Paired- end	MSS(WGS)	Comment
	SOAPdenovo	٠	1.05			V		
	ABySS 2	 Image: A state of the state of	1.3.2			V		Maximum K-mer value is 64.
0	Velvet 🗗	۲	1.2.03			~	~	We severe recommend when performing Velvet, total length of those reads is up to 22G bp.Maximum K-mer value is 64.
2	Trinity @	۲	r2012- 06-08			~		RNA-Seq De novo Assembly

最下部の「NEXT」をクリック。

Trinityの実行 クエリのレイアウト選択

実行するAccessionの横のチェックボックスをクリック

右側の「confirm」ボタンをクリック。(ペアエンドのクエリの場合「Set as PairEnd」ボタン)

Layout of s	ingle sequence.					
5'			3'			
Linker(1)	Target	Linke	er(2)			
	-			T		
	Run ACCESSION	Read length	Quality Score			
5509	_SRR042533_e.fastq.bz2	bp				

画面下に確定したレイアウトが表示されるので、最下部の「NEXT」をクリック。

PairedOrientation	RunAccession	RunAlias	RowLength	QualityScore1	QualityScore2	
single	1819	SRR042533 by Preprocessing				

今回はクエリファイルを1つしか選択していないので、あまり意味はないが、 複数のファイルを選択していた場合、それらをすべて結合して実行するか、 あるいは、別々に連続して実行するかをこの画面で選択する。

Trinityの実行 実行オプションの指定

library type および 実行時のオプションを指定。 今回はデフォルトの条件で実行する、

end analysis	
and-Specific (Forward) OStrand-Specific	c (Reverse)
ceMax 4GbflyGCThreads 1CPU 4	
length 201	501に変えてみましょう
fa : for fasta file]	
from pileupped reads to <u>submit WGS divi</u>	sion of DDBJ.
	and-Specific (Forward) Strand-Specific ceMax 4GbflyGCThreads 1CPU 4 length 201 fa : for fasta file] from pileupped reads to <u>submit WGS divi</u>

参考) Pipelineで使用している Trinity 実行コマンド

クエリファイルの種類 FASTA or FASTQ (自動で指定される) メモリ、CPL Trinity.plseqType fqJM 100GbflyHeapSpaceMa	」関係の指定(固定) x 4GbflyGCThreads 1CPU 4
single <クエリファイル名>output <出力ディレクトリ名>	min_contig_length 201
入力ファイル・出力ファイルの指定 (自動で指定される)	ユーザーの指定するオプション

Trinityの実行 実行オプションの確認

メールアドレスを入力して、「RUN」ボタンを押す

inen die request is	completed, the s	ystem sends an email to this a	ddress.			
hnagasak@lab.nig	1.ac.jp		* R	equired		
Result files will be de	leted 60 days afte	r submission.				
Assembly [trinity]						
Query sets						
Query set1						
PairedOrientation	RunAccession	RunAlias	RowLength	QualityScore1	QualityScore2	
single	1819	SRR042533 by Preprocessing		l		
Assembly comma	nds					
trinity						
Set optional	parameters	of the single-end analy	vsis			
Step1) Assemb	ly					
Specify the library	type : 💿 not Stran	nd-Specific 🔵 Strand-Specific (F	orward) St	trand-Specific (R	everse)	
Trinity.plseqType	e fq(or fa)JM 10	0GbflyHeapSpaceMax 4Gbfl	lyGCThreads	1CPU 4		
single reads.fq	output output_dir	min_contig_length 201				
seqType is autom	atically selected.	I fo : for fasto file, fa : for fasta file	e 1			
and the second						
	ssembled seque	nces in FASTA file from plieupp	bed reads to s	ubmit WGS divis	ion of DDBJ.	
Step2) Create a						
Step2) Create a Set filtered length	for contigs					

Trinityの実行 実行状況の確認

Status \rightarrow denovo Assembly から、実行したジョブの確認をする

	ID	UserID	Submission accession	P/S	Status	Tool	Read #	Read length	Assembly detail	Mapping detail	Start time End time	Elapsed time
0	5516		 Whole transcrip	S	running	Trinity	46,765,342				2013-04-30 18:26:32	
•	5515	koshu01	SRA012701 GSM497271_1	S	complete	Trinity	7,468,448	(View)	2013-04-30 18:15:21 2013-04-30	02:40:24
0	5514	koshu01	 SRR042533 by	S	complete	Trinity	7,420,316		View		20:55:45 2013-04-30 18:13:59 2013-04-30	02:42:20
	5508		Drosophila RNA	S	complete	Trinity	18,524,700				20:56:20 2013-04-30 17:23:42 2013-04-30	03:43:55
0	5507		SRA009364 42CRDAAXX	P	complete	Trinity	9,262,350				21:07:38 2013-04-30 15:45:56 2013-04-30 21:07:35	05:21:38

「View」ボタンをクリックして、詳細確認。

Trinityの実行 実行状況の確認

Status \rightarrow denovo Assembly から、実行したジョブの確認をする

ID								
5514								
Tool (Version)								
Trinity (r2012-06-08))							
RunAccession or Fil	ename	Download		R	ead length	Alias		
1819		5509 SRR04253	3 e.fastq.bz2		N.A. bp	SRR04	2533 by Preprocess	ing
Download modified o	lueries							
• 5509 SRR04	2533 e.fastq.gz (Ori	ginal size 1.3 GB	2					
Download wgs file								
• out WGS.fast	a.gz (Original size 1.	0.MB)						
Assembly statistics								
			結果フ	7ァイル	の統計値		Contig # Total contig size : Maximum contig si Minimum contig N50 contig si	: 2,466 1,018,683 ize : 4,351 size : 202 ze : 450
Time								
		Start time				1	End time	
Wait time	2013-04-30 18:1	3:59		20	13-04-30 20):56:20		
Wait time 0: 0:47						結果	ファイルのダワ	ウンロ-
Wait time 0: 0:47			Otra di stati	End tim	e Log1	Log2	Result	MD
Wait time 0: 0:47	Command		Start time					

「BACK」ボタンで、一覧画面に戻る これで基礎部は終了です。

DDBJパイプライン高次解析部による RNA-Seqアセンブル結果の解析

高次解析部起動

パイプライン基礎部の左のメニューカラムから「step-2/Workflow」を クリック。





高次解析部(GALAXY)が起動

Tips:

http://p-galaxy.ddbj.nig.ac.jpでURL 直打ちして、「ツール」メニューの 「Work Flow」をクリック。 基礎解析と同じパイプライン登録時の メールアドレスとパスワードを入力し ても起動可能。



analytical pla	Annotation Pipeline is a cloud-compu- tform for next-generation sequencing	data.
User ID:		
Password:		
	Login	

RNA-Seqのアセンブル結果をインポート

TrinityによるRNA-Seqのアセンブル結果を GALAXYにインポートする。

左側「ツール」メニューの「Work Flow」をクリック

左側「ツール」メニューの「COMMN PROCESS」 の下「import contig form DDBJ Pipeline」を クリック

実行したジョブのsamfileのリストのうち、今回は 「SRR042533 by Preprocessing」の「import」を クリック







中央にツール実行開始の表示が現れ…

左側のヒストリーに読み込み中のファイルが 表示される(緑色になったら終了) ヒストリーの目のアイコンをクリックすると 中央にプレビューされる。



さらにその下の「transcriptsToOrfs (N.A.) Trinity Transcripts to Candidate Peptides」 をクリック

CPU: 16くらい推奨 「Execute」をクリック

結果としては 1) アミノ酸配列 2) pfamのドメインとのマッチング 3) その他ORF候補 が返ってくる。



0090	結果1
.A.) Trinity date	<pre>>m.565 g.565 ORF g.565 m.565 type:internal len:207 (-) DLEMGIEGLKEELIFLKKNHEEELLANRAQMSGQVHVVERAPAEDLTKVHADIREHVES ITAKNQKELETYFNSKSBALIKENHTQTVTLQTFSREVTEVKKSLQALGIELESLLGMKA SLEGTLQDTONRYSMHLAGYQQQVTSLEQQLVQLRADLVRQGQDYQMLLDIKTRLELEIA EYRRLLEGEAAASSSTSSTSTKTRRL >m.566 g.566 ORF g.566 m.566 type:complete len:216 (+) MAQSVFVVMFKLVLVGDGGTGKTTFVKRHLTGEFEKKVATLGVEVHELFPNTNRGNVKF NVMDTAQGEFGGLROGYTJQACAIIHEPTSKVYKNYMFMRDLVKVCEHIFVLCG</pre>
ion 0.0.1)	NKVDIKDRKVKAKSIVFHRKKNLQYYDISAKSNYNFEKPFLWLARKLIGDPNLEFVEMPA LAPPEVTMDPALAVQYEKELHVASQTALEDDEDDL*
astA format: a 1 : n amino acids): both strands :	>>m.568 g.568 g.568 g.568 m.568 type:internal len;227 (-) GDRFREDRKRLPEKSIDHILIDDODPNSCBSRIPUIGEWNKAAIGGQNSEFSLGFGN DVKYPFLDVMSRENNGLARRIYEGSDAALQLQGYDEVSSPLLLDVDLRYPDNAVDSLTT NQFSQLPNGSEIVVAGRLKDNDIDNEPVEVFQGLNDFSEGQGFSVLDWSGMYPDDDYIF GDTTRLHAVIITIQQLLDKSKTGDAEEKANASAEALDMSLRYSFVTP >m.571 g.571 g.571 m.571 type:5prime_partial len:394 ASGGGCHTSGCSWFNAGAKOPFSVPYSIDPNYKCKTSGEEISSYDDUGVORCRLVS LLDLALEKDYVRGKVADYMNRLVDWGVAGFRVDACKHWWFGDLSAVYGRLMNLNTKWFPE GSAPFIPGEVIDLGGALSITYVVHLGRAGFRVTPAKIALERYGTWKWSSFR NRHIVNGKDQNDWMGPFSUFDGSTKSVPINPDETCGDGWVCEHRWRQIKNWVIFRNVVNG QPHSNWNDNNSQVAFGGNARFIIFNDDAUDVLNTGLARGYCLVISGQKEAGCCT GKQIHVGSGGRAHFRISNRDEDPFVALHVESKL*
い推奨 mmscan	>m.573 g.573 g.573 g.573 g.573 m.573 type:Sprime_partial len:224 WEPSWP0VSLQETGFHECGSLINENVUTALACONTENDENENIOV MQVGQVFKHPNINSYTINDITLIKLASPAQLNIRVSPVCVAETSDVFPGGMKCVTSGWG LTRYNADPTPELQQVALFLTHEECKRHWRGKITDLWCCAGSGASCSGGDGGGGLVCE KAGAWTLVGIVSWGSGFCSVSSPGVYARVTMLRAMMDQIIAAN*

結果2	#			full s	equence		best 1	domain		do	main	numb	er e	estim	ation			Î
	# target name	accession query name	accession	E-value	score	bias	E-value	score	bias	exp	reg c	lu	ov e	env d	om re	p inc	description of target	
	#Actin	PF00022.14 m.1	-	2.8e-162	539.5	0.0	3.2e-162	539.3	0.0	1.0	1	0	0	1	1	1 1	Actin	
	Apolipoprotein	PF01442.13 m.3	-	1.1e-38	132.6	10.6	1.1e-38	132.6	7.3	1.8	2	0	0	2	2	2 2	Apolipoprotein A1/A4/E	
	domain																	

結果3	comp1002_c0_seq10	621	ID=m.565;Name=ORF_g.565_m.565_type:internal_len:207_(-)_(g.565,_m.565);	0	-	0	621	1	621	0
	comp1006_c0_seq137	685	ID=m.566;Name=ORF_g.566_m.566_type:complete_len:216_(+)_(g.566,_m.566);	0	+	37	685	1	648	0
	comp1010_c0_seq12	683	ID=m.568;Name=ORF_g.568_m.568_type:internal_len:227_(-)_(g.568,_m.568);	0	-	2	683	1	681	0

RNA-Seq由来のアミノ酸配列をBLASTPにかける



ワークフローの保存も可能

(GALAXYがメールアドレスを訊いてきたりするのでパイプラインのユーザーアカウント取得後)

参考: https://main.g2.bx.psu.edu/u/aun1/p/galaxy101 の"4. Converting histories into workflows"など



参考資料

DDBJパイプライン(基礎部)へのアカウント作成

DDBJパイプライン(基礎部)に新規登録

http://p.ddbj.nig.ac.jp/

DDBJパイプライン(http://p.ddbj.nig.ac.jp/)に入る。

「New account」をクリック



Note that this account is As DDRJ Pipeline is a w	NOT registered as a NIG supercomputer account
As DDBJ Pipeline is a w	no registre a a mo supercompare account.
here / Supercomuter II	abservice of NIG supercomputer, user information was publicly opened to the internet from
After registration your	III receive a confirmation email with your user ID and initial nassword. Please input your
email address correctly	
" UserID:	Use 6 to 16 charactors.
* Emsil addrage:	
cillan audress.	
* Retype email	2
address:	* for confirmation.
-	
- First name:	
* Last name:	
* Institution	
with	Contester Information Biology, National Institute of Constice

Registration

UserIDを決めて必要情報を入力

「Registration」をクリック パスワードがeメールで届くのでそのパスワードでログイン



Query file指定方法 FTP Upload画面へ遷移

ACCOUNT	(Hui	ning .	Status								
login ID (guest)	Se	elec	ting Qu	ery Files							
ANALYSIS Data setup											NEX
DRA Start	FTF	p ubl	oad Priva	ite DRA ent	ry O Import	public DRA	HTTP up	load			
HTTP upload	N	letada	ita of the DRA	entry.		- 入力	יעכנ	イルの	の指	定方法	₹4種
		11.						Sele	ect a meta	adata : DRAO	00001
.メニューの	-12	Up	Dioad a	2793	19		FILENA	ME	1	DL	VIEW
step-2	Sub	missio	n DRA000001				DRA0000	001.submiss	ion.xml	(DownLoad)	View
Genome (SNP/Short	Sam	ple	DRS000001	Bacillus subt plasmid pBE	lis subsp. natto f ST195L	3EST195 without	DRA0000	001.sample.	xml	(DownLoad)	View
RNA-seq (Tag count)	Stud	y	DRP000001	Natto BEST1	95		DRA0000	001.study.xm	ıl	DownLoad	View
Genome	Expe	erimer	t DRX000001	NATTO_BES	T195_SEP08		DRA0000	001.experim	ent.xml	(DownLoad)	View
(Large Indel)	Run		DRR000001	2008-09-12.8	BEST195-Lane7	1	DRA0000	001.run.xml		(DownLoad)	(View
Job Confirmation	1.0000					2	and a second sec		-		
step1.	STUDY TITLE			Whole genome sequencing of Baillus subtilis subsp. natto BEST195							
step1.	STU	JDYT	YPE	Whole Genome	Sequencing		la antin'ny taona 2014. N	MARK.			
step 1. Assembly status	S	elect	your registere	i query files.							
step2-All status	Differ	entin	strument model	s can't he select	ed together						
Help											
HELP @	Cani	gie	pared and	Cample	0				Deed	to a to see a set	
MANUAL		No.	ACCESSION	ACCESSION	ACCESSION	STRAIN	Run_date	Read #	length	model	Layou
									-		



Query file指定方法 FTP clientによるUpload

pioad FASTA	VFASTQ f	iles Select a FASTA/F	ASTQ file Input a specification	
Please u	pload	the file to be use	ed.	
For security FTPS.	this ftp se	erver is using FTP over S	SL protocol. Therefore, please use FTP client that supports the file transfer protocol	
Server : Port		p.ddbj.nig.ac.jp:21	1.FTP clientをローカルPCにインスト-	ールし、
User ID/pa	User ID/password Your Pipeline login ID/password		assword DDBJのサーバーへFTP転送をする。	
FTP clie	nt softwa	re.	*転送方法は次ページに記述	
Windows	WinSCP,	FileZilla		
MacOSX	FileZilla,	Cyberduck etc		
The unload	directory i	s not open to the other u	users. 2.デー タ転送	É後、次へ
The upload		al huu aal		



FTP client Cyberduckの場合	
1.http://cyberduck.ch/ヘア	クセス の の 日本派
Cyberduck オープンソース のFTP、SFTP、WebDAV、Cloud Files、 Google Docs、Amazon S3用ブラウザ、MacとWindowsに対応。	
Cyberduckについて ニュ	ユース 更新複歴 開発 ヘルプ 客付 2.ダワンロード
made シビードコンドローク2 (Mice Ovinde O Sample) ドンドのランジロード a m イキ www.iterativewer.com FTP Site - Free Trial 1,515,353+ Business Users Trust Our Award-Winning Solution. Free Trial! Unleaded ShareFie.com MacOS X Snow Leopard 登場さらに進化 より高速に もっと簡単に 世界で最も先進的なOS。今ずぐ購入 store apple Ade by Google	ダウンロード バージョン3.8.1 2010年12月6日 Cyberdyek3-8.1 zip
あらゆるサーバに接続。 使いやすいインターフェースで、ETP (File Transfer Protocol)、ETP/TLS (FTP secured over SSL/TL5)、SETP (SSH Secure File Transfer)、WebDAV (Web-based Distributed Authoring and Versioning)、Amazon 53、Google Storage、Google Docs、 Windows Arure、Rackspace Cloud Filesに接続。Eucalyptus や OpenStackのオープンな最新クラウドソフトウェアも利用でき ます。	ユニバーサルバマイナリ、Mcc OS × 10.5以降が必要 バージョン4.0パブリックペータ 2010年12月13日 <u>Cyberdusch-Installer-4.08&exe</u> Windows XP、Windows Yato または Windows 7か必要
Open Connection Quick Connect Action Edits Refresh Disconnect	Downloads hosted by <u>Carteboy CDN Gen</u> <u>Source Cantent Delwry</u> <u>寄付</u>
(I) () (cyberduck.ch) () ()	Vita 🛄 530
Lookmark O History 😵 Bonjour	
Amazon S3 (HTTPS) si amazonaws.com https://docsars/scom/com/codwrdark.ch	



Query file指定方法 通信先サーバ情報を設定





Query file指定方法 Upload

\varTheta 🔿 🔿 👔 👔 🙆 🙆	SL Get a donation key!	
) 🧷 🚔 🗁	ストデータ:
新規技術 ジイック技術 アクジョン 史和	suk	omission DRA000001
1 基礎部なのでQueryフォ	ォルダをダブル sar	nple Bacillus subtilis subsp. natto BEST195
▶ galaxy	WIL	nout plasmid pdes i 1952
▶ auery クリック	Re	ad数 : 9,977,388
	Re	ad length : 36
🙆 🔿 🧿 🔹 133.39.116.60 – FT	P-SSL Get a donation keyl	
		コーノリナッドニッグのドロップ"ナス
新規接続 クイック接続 アクション 引	2.0pload 072い7	(アイルをトラック&トロック 9 る。
(m) (+) (/query	; A Q	
ファイル名 ▲ サイズ	変更日	
		33.39.116.60 - FTP-SSL
		🖸 🏟 🧑 🥢 🛛 🚔
	新規接続 クイック接続	アクション 更新 編集 接続解除
	📖 🕨 📄 🚺 /query	• A Q
	ファイル名	▲ サイズ 変更日 4.3 CB 今日 8:16
	test1_2.fastq	4.3 GB 今日 8:14
		3.Upload完了⇒Pipelineの画面へ
	2 ファイル	24



DDBJ

DNA Data Bank of Japan

Query file指定方法

Uploadしたファイルの注釈づけ1

1.Pipelineの画面に戻る						
UploadしたファイルがSingle-endの場合	UploadしたファイルがPaired-endの場合					
2.Select a FASTA/FASTQ fileを選択	2.Select a FASTA/FASTQ fileを選択					
Image: Second FASTA/FASTO file Second FASTA/FASTO file Second FASTA/FASTO file Please specify read layout and correspondent files, Image: Second FASTA/FASTO file Image: Second FASTA/FASTO file Second FASTA/FASTO file 3.Single-end & Eff. Image: Second FASTA/FASTO file Image: Second FASTA/FASTO file 3.Single-end & Eff. Image: Second FASTA/FASTO file Image: Second FASTA/FASTO file 3.Single-end & Eff. Image: Second FASTA/FASTO file Image: Second FASTA/FASTO file 3.Single-end & Eff. Image: Second FASTA/FASTO file Image: Second FASTA/FASTO file 3.Single-end & Eff. Image: Second FASTA/FASTO file Image: Second FASTA/FASTO file 3.Single-end & Eff. Image: Second FASTA/FASTO file Image: Second FASTA/FASTO file 3.Single-end & Eff. 3.Single-end & Single Image: Second FASTA/FASTO file 3.Single-end & Single 3.Single-end & Single Image: Second FASTA/FASTO file 3.Single-end & Single 3.Single-end & Single Image: Second FASTA/FASTO file 3.Single-end & Single 3.Single-end & Single Image: Second FASTA/FASTO file 3.Single-end & Single 3.Single-end & Single Image: Second FASTA/FASTO file 3.Single-end	Registration of fastq/fasta files Upload FASTA/FASTO % Select a FASTA/FASTO % Please specify read layout and correspondent files. Please specify read layout and correspondent files. Read ispat: Paired-end Select a FASTA/FASTO file: Imename Not select Not select Select a reverse Type Not select Select a reverse FASTQ file: Aread1 fileを選択 Hype All Paired-end query, please select a "Reverse" query file. Select a reverse FASTQ file: Imename Ype Select a reverse FASTQ file: Select a reverse FASTQ file: Imename Not select Select a reverse FASTQ file: Select a reverse FASTQ file: Select a reverse FASTQ file Select a reverse FASTQ file Select a reverse FASTQ file Select a reverse FASTQ file <t< th=""></t<>					

Query file指定方法 Uploadしたファイルの注釈づけ 2

Registration of fastq/fasta files
Upload FASTA/FASTQ files Select a FASTA/FASTQ file Input a specification
Please specify instrument model.
SelectedFile 1 test1_1.fastq
SelectedFile 2 test1_2.fastq 1シークエンサの継話を認む
Read layout Paired-end 1.シークエンリの依律を迭が
(Required) Study title title と、STUDY TITLEを入力
NOTICE: After confirming your entries, push the SUBMIT button to register uploaded files.
SUBMIT
3.
Registration complete.
Press "Mapping / Assembly" button, to goto job input pages.
Assembly / Mapping
4.Assembly/Mappingの実行画面へ



Query file指定方法

Uploadしたファイルの確認

Uploadしたファイルを使用して解析が可能になって いる。

TP upload Private DRA entry List of your uploaded files by FTP client. [A	A/FASTQ(FTF Import public DF	P client)を	·選択 upload	NEXT
Filename	Description	Layout	Instrument model	File size
GSM727564_d0Foxh1.bed.gz	Foxh1	single	ILLUMINA	124.4 KB
test1_1.fastq (more 1 files)	test data	paired	ILLUMINA	3.4 GB
		_		DELETE NEXT