

メニーコアプロセッサ PEZY-SC3 による ヒトゲノム解析の高速化と PEZY-SCシリーズの展望について

Supercomputing Japan 2025

2025/02/04

PEZY Computing 酒井智史

sakai@pezy.co.jp

Contents

- メニーコアプロセッサPEZY-SC3とZettaScaler3.0サーバーについて
- 高速ヒトゲノム解析システムZettaVEGAについて
- ZettaVEGAによる大規模ヒトゲノム解析
- PEZY-SCシリーズのロードマップ

PEZY Computingについて



- ・プロセッサ、PCBボード、システム、ソフトウェアの開発と販売
- ・超音波診断装置、組込み制御システムの開発・製造・販売
- ・液浸冷却装置、液浸冷却方式によるコンピュータシステム、液浸冷却用のシステムボード等の開発・製造・販売
- ・ゲノム解析システム、医療画像表示システム、画像解析AIシステム、ゲノム関連創薬システムの開発
- ・人工知能チップ、アルゴリズムの開発

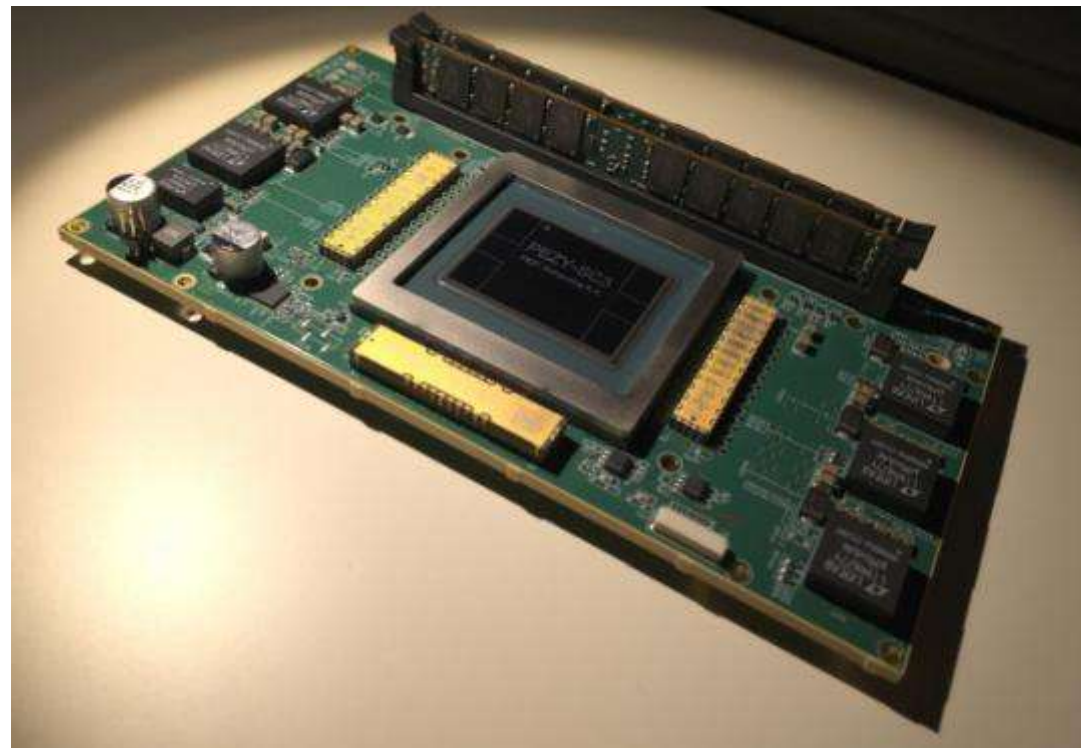
2017年11月 暁光 Top500 4位, Green500 5位

会社HP: <https://www.pezy.co.jp/>

PEZY-SC3 の仕様

演算コア(PE)数: 4,096
動作周波数: 1.2 GHz
最大理論性能: 19.66 TFLOPS(DP) /
39.32 TFLOPS(SP) /
78.64 TFLOPS (HP)
搭載メモリ: HBM2 4ch, 32GB
メモリ帯域: 1200 GB/s
インタフェース: PCI Express Gen 4 × 16
プロセスノード: TSMC N7
消費電力: 600 W

OpenCLライクなPZCLにてプログラミング可能です



ZettaScaler 3.0 Server Unit の仕様

- **ZettaScaler 3.0 Server Unit仕様**

- ベースボード: ExaScaler EPX-BASE2 x 1
- CPU: AMD EPYC (64 core) x 1
- アクセラレータ: PEZY Computing MOD-SC3H (PCIe Gen4x16 bus) x max 4 modules
- メインメモリ: DDR4 ECC Registered 3200MHz SDRAM max 2TB (128GB DIMM x 16) ※標準は1TB
- ストレージ: M.2 NVMe SSD 2TBx4
- OS: AlmaLinux9・Ubuntu



※ 構成については柔軟に対応可能な部分もございます。
お気軽にお尋ねください。

ZettaScaler 3.0の実績

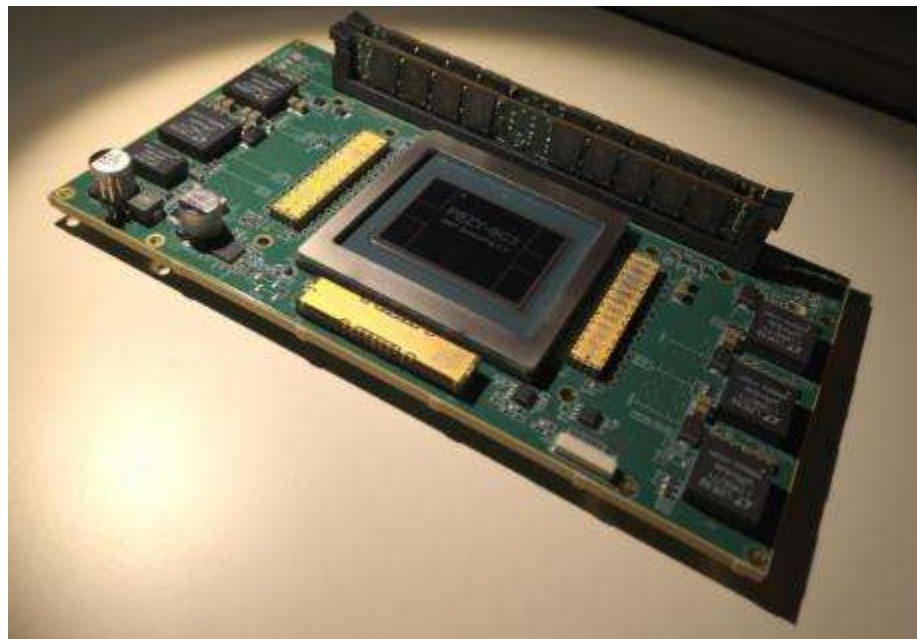
2021/11 NA-IT1 (NA Simulation) Rmax 1,684.83 [TFlop/s] , Power Efficiency 24.682 [GFlops/Watts],
Green500 12位, Top500 435位

2022/06 NA-J2 (NA Simulation) Rmax 1,952.46 [TFlop/s], Power Efficiency 21.65 [GFlops/Watts],
Green500 27位, Top500 401位

2022/11 NA-J2 (NA Simulation) Green500 34位, Top500 438位

2023/06 NA-J2 (NA Simulation) Green500 39位, Top500 481位

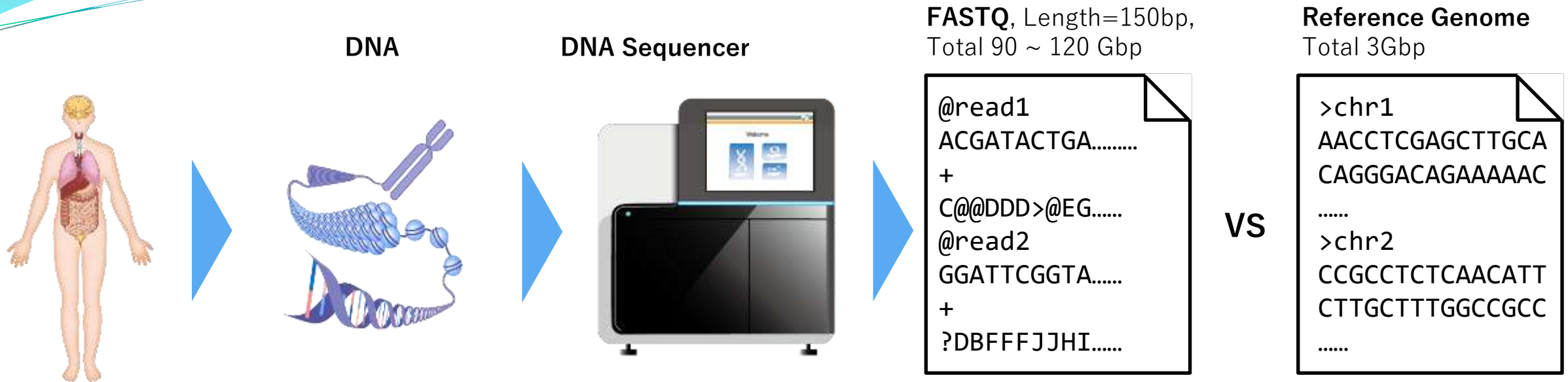
※上記実績は空冷サーバーではなく液浸タイプのサーバーを使用したものです



Contents

- メニーコアプロセッサPEZY-SC3とZettaScaler3.0サーバーについて
- 高速ヒトゲノム解析システムZettaVEGAについて
- ZettaVEGAによる大規模ヒトゲノム解析
- PEZY-SCシリーズのロードマップ

ヒトゲノム解析について

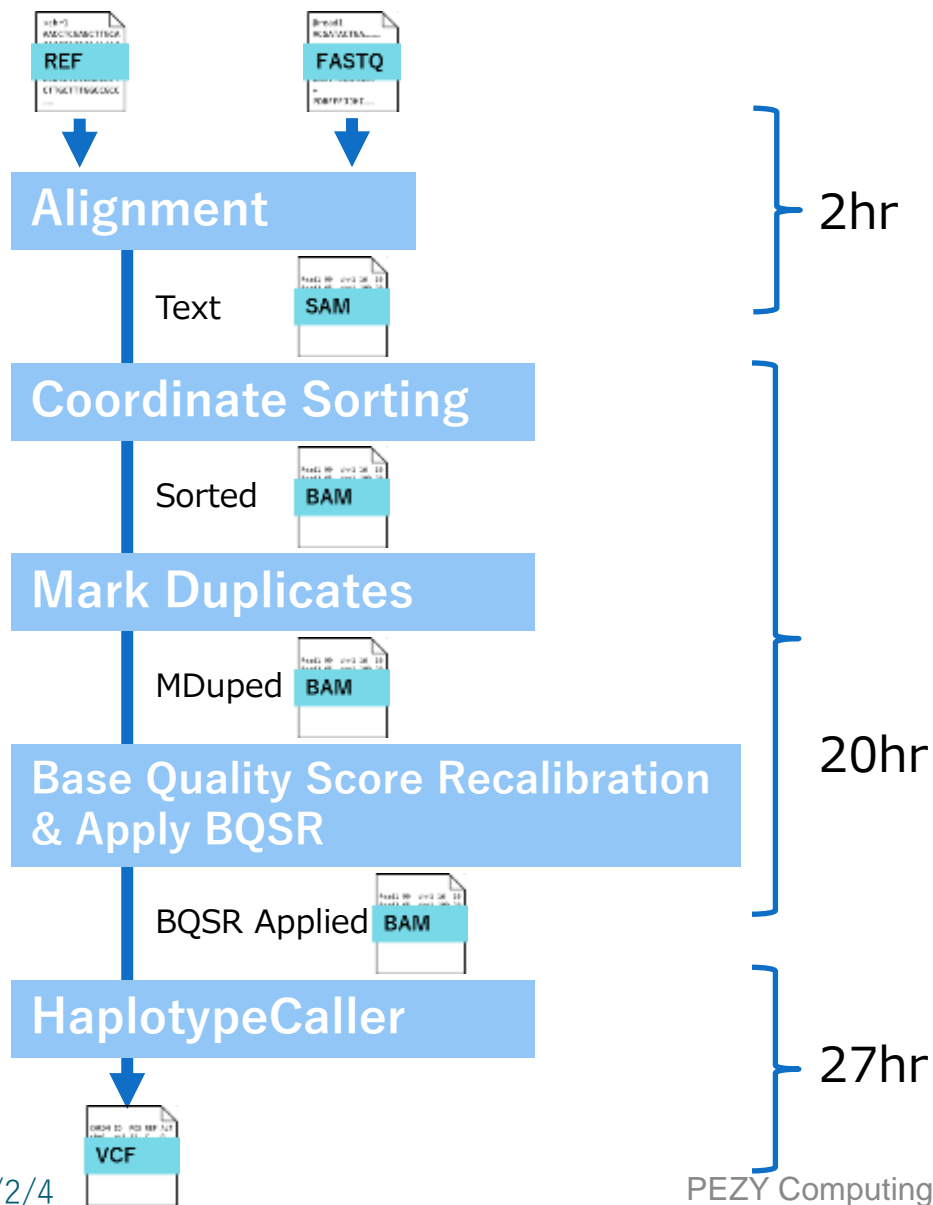


出力: リファレンスゲノムと個人のゲノムの差 (VCFファイル)

ヒトゲノム: ATCG四種類の文字で表される約**3Gbp** (base pair) の文字列
基準となるヒトゲノムの文字列**リファレンスゲノム**を解析に使用
リファレンスゲノムにはGRCh37/HG19, GRCh38/HG38, T2T-CHM13などがある

入力: 150bpの細切れの文字列 (read) が合計で**90 ~ 120 Gbp** (FASTQファイル)
高スループットなDNA シーケンサは短いDNAの文字列を超並列で解析
入力ファイルにはDNAの文字列以外に、IDや解析の信頼性の値も含まれる

GATK (Genome Analysis Tool Kit) Best Practice



Alignment:

150bpの細切れの文字列をクエリとし、リファレンスゲノム中から最も一致する箇所を検出し整列させる

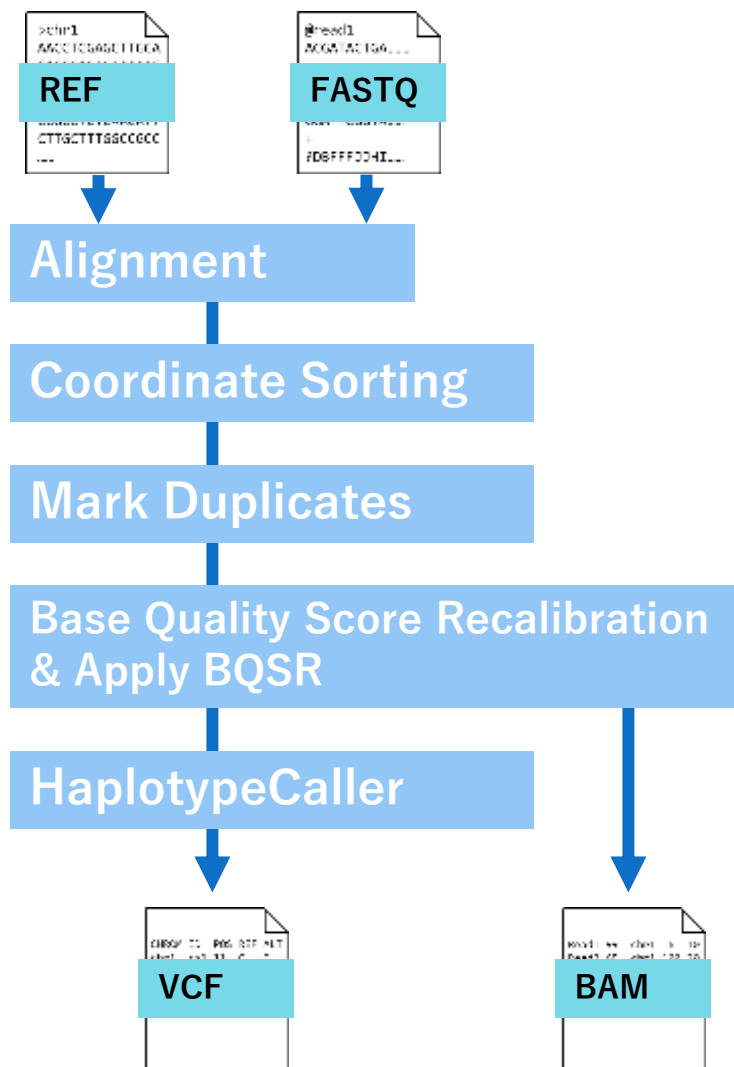
アライメント結果の後処理:

Coordinate Sorting, Mark Duplicates, BQSR
バイナリ形式への変換、重複リードへのFLAG付与、塩基配列の信頼性の値の補正を行う

HaplotypeCaller:

バリエーションコール
アライメント結果とリファレンスゲノムを比較し、リファレンスゲノムと個人のゲノムの差を出力する

ZettaVEGAの概要



ZettaVEGA : ZettaScaler Verifying Environment for Genome Analysis はZettaScalerで動作するヒトゲノム解析パイプラインです。

ヒトゲノム解析において最もよく使用されている、**GATK Best Practice** のパイプラインで使用されているソフトウェアを高速化し、100GbaseのFASTQを入力としたときに、HG19/GRCh37リファレンスを用いた時には**15分以内**で、HG38/GRCh38リファレンスを用いた時には**25分以内**で、バリエーションコールまで完了します。

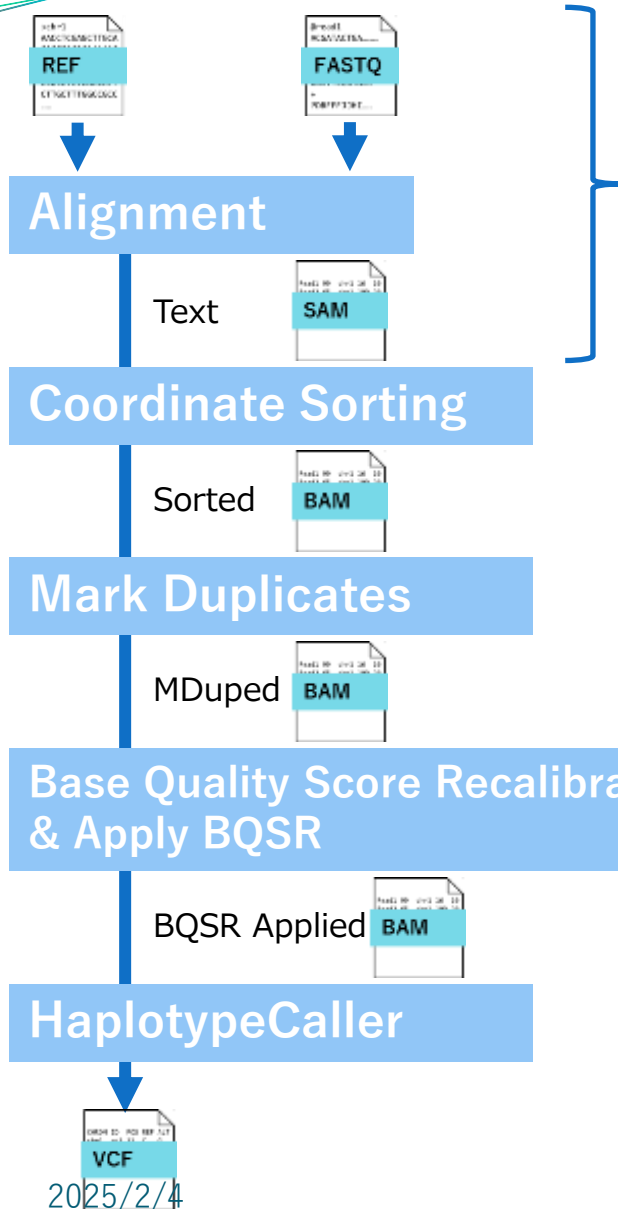
高速化のために以下の改良を行っています

- PEZY-SC3のメニーコアを活かした超並列処理
- 中間ファイルをオンメモリ化
- CPU処理の最適化（パイプライン処理・並列化）

高速化だけでなく精度の向上も行っています。

- BWA MEMに精度向上のためのオプションを実装
- GATK 4.2 で実装された精度の高い確率モデルを実装

Alignment



処理内容

FASTQファイルの各Readのリファレンスゲノムの位置を割り当てる。
ただし、ゲノムの個人差やDNAシーケンサのエラーに由来する文字の置換・欠失・挿入がある。

リファレンスゲノム

.....ATTGCGCCCTTGAAATTAAGGAAAAGAAATTTTCG.....

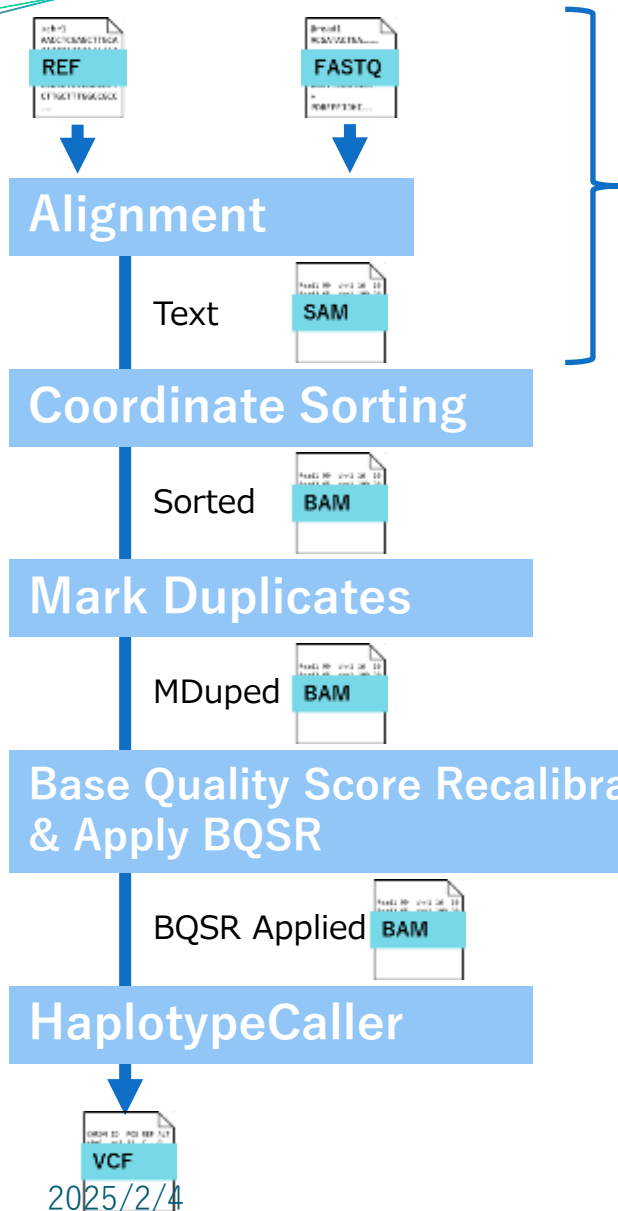
Read1

CGCCCTAGAAA

Read2

AAAGAAATTCG

Alignment



処理内容

FASTQファイルの各Readのリファレンスゲノムの位置を割り当てる。
ただし、ゲノムの個人差やDNAシーケンサのエラーに由来する文字の置換・欠失・挿入がある。

リファレンスゲノム

.....ATTGCGCCCTTGAAATTAAGGAAAAGAAATTTTCG.....
CGCCCTAGAAA AAAGAAATT-CG

Read1
CGCCCTAGAAA

Read2
AAAGAAATTCG

Alignment: BWA-MEMの計算

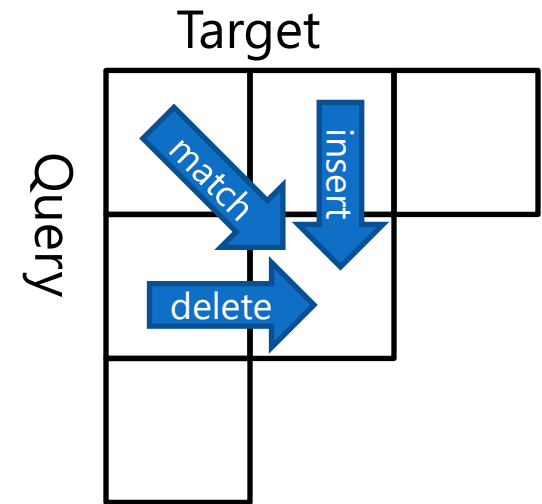
Smith-Waterman Alignment (SWA)

- 動的計画法によって重み付きの編集距離を求める
- リファレンスとの一致度を求める計算(アラインメント)に使う
- $n = \text{length}(\text{Target})$, $m = \text{length}(\text{Query})$ としたとき、計算量は $O(nm)$
- 演算律速

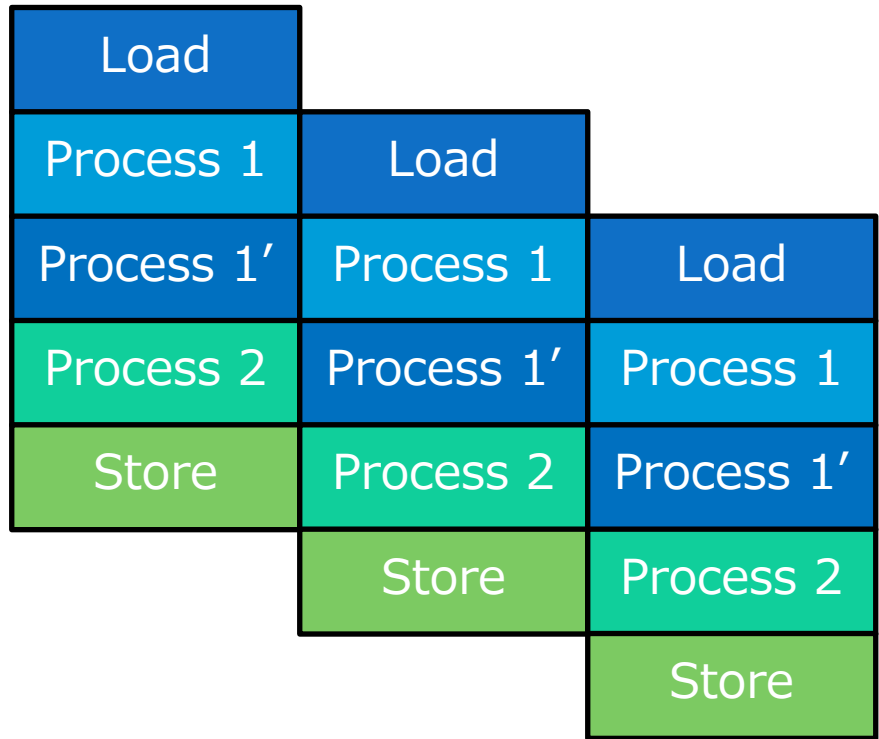
FM-index

- Burrows-Wheeler変換と接尾辞配列を用いた検索アルゴリズム
- “seed”の検索に使う
- $n = \text{length}(\text{Target})$, $m = \text{length}(\text{Query})$ としたとき、計算量は $O(m)$
- メモリアクセスが律速になる

BWA-MEMはFM-indexを用いて完全一致する最長配列を検出してSeedとし、Seedを起点にSWAを行い文字列の一致度を計算する。



PEZY-SC3を用いたBWA MEMの高速化



多数の短いリードが入力となるため
PEZY-SC3でリード単位での超並列処理

FM index + SWA

process1 の残りをCPUで処理

ペア情報を使って補正
主にCPU処理。オプションでSWAをSC3にオフロード。

work imbalance対策

- 一部のリードが極端にメモリ・時間のリソースを消費する
- 処理に一定以上のリソースを必要とする入力に対しては、SC3での処理を打ち切ってCPUでやりなおす(process1')

Sort, MarkDuplicates, BQSR, Apply BQSR

Alignment

Text



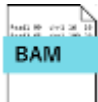
Coordinate Sorting

Sorted



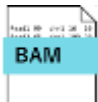
Mark Duplicates

MDuped



Base Quality Score Recalibration & Apply BQSR

BQSR Applied



HaplotypeCaller

処理内容

アライメント結果のソート・重複チェック・クオリティ値補正

実装と高速化

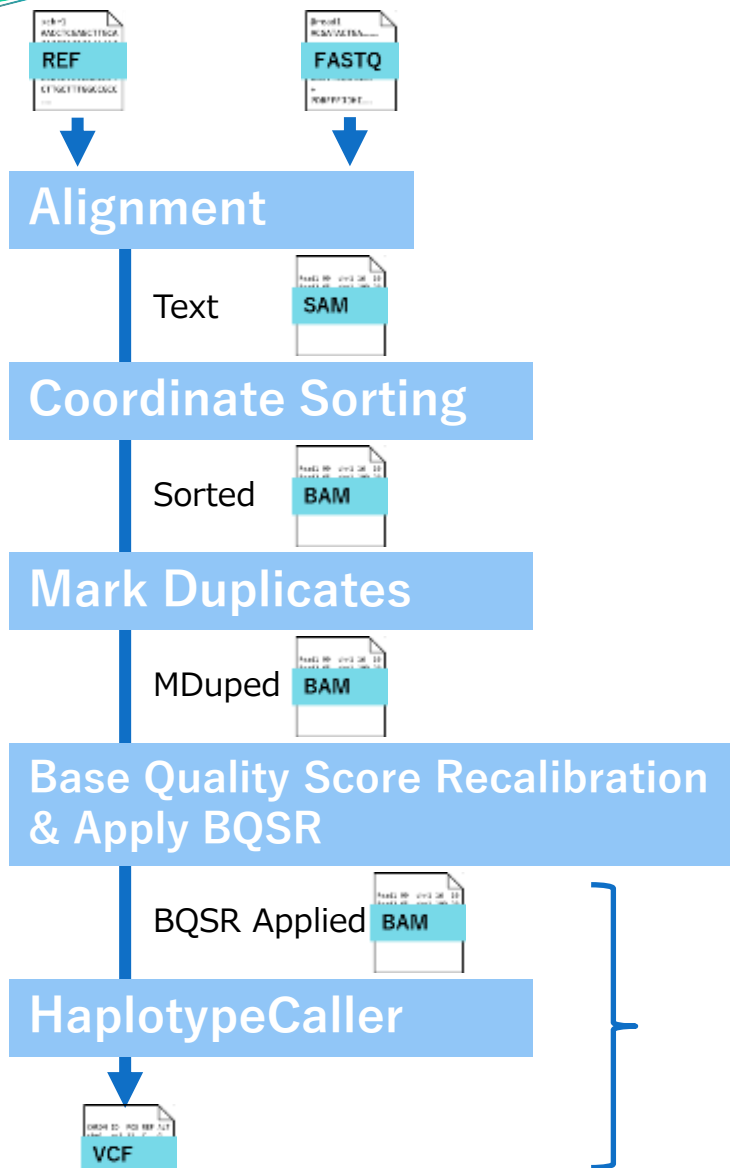
- ・オンメモリ化

元実装はメモリ節約のため多数の中間ファイルを作っていた
メモリが1TBあれば200Gbp程度あっても余裕をもって載せられる
BWA-MEM、HaplotypeCallerとは共有メモリを介して通信
BWA-MEMはテキストデータを經由せず

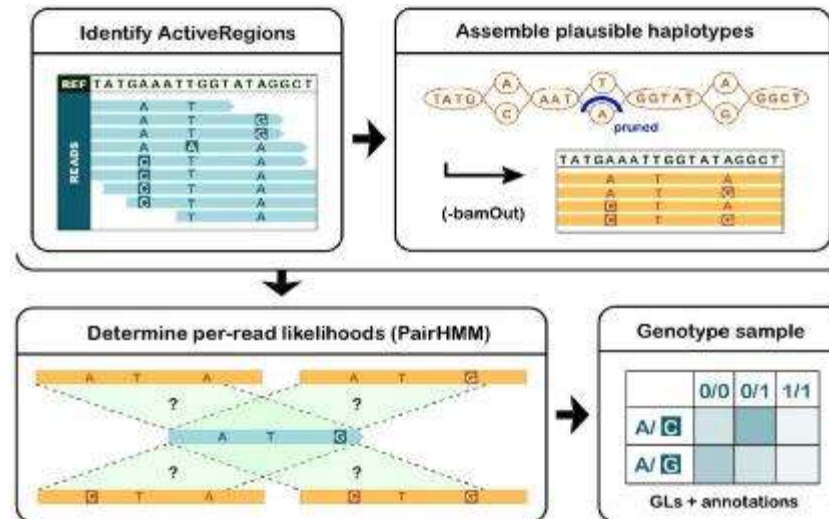
- ・シンプルな処理なので、C++で自前で実装&並列化

・CPUでも十分に速く、また通信のコストの方がメリットを上回る見込みのため、SC3デバイスへのオフロードは行わなかった

HaplotypeCallerによるバリエーションコール



1. リードを先頭から順番に読み込む
2. 変異がありそうな領域(Active領域)を切り出す
3. Active領域内リードのdeBruijnグラフから配列候補を生成
4. 配列候補からリードが生成される尤度を計算 (PairHMM)
5. SWAによる配列候補vsリードのアライメント
6. 変異が存在する尤度を計算
7. 閾値以上なら各種統計情報とともに出力



Pair-Hidden-Markov-Model
アライメント計算の一種
計算量 $O(nm)$

R. Poplin, et al., doi:10.1101/201178

HaplotypeCallerの高速化

1. リードを先頭から順番に読み込む
2. 変異がありそうな領域(Active領域)を切り出す
3. Active領域内リードのdeBruijnグラフから配列候補を生成
4. 配列候補からリードが生成される尤度を計算 (PairHMM法)
5. SWAによる配列候補vsリードのアライメント
6. 変異が存在する尤度を計算
7. 閾値以上なら各種統計情報とともに出力

順番に行う必要があり
パイプラインで隠蔽

↓
ここから先は
領域毎に並列化

↑
比較的単純かつ大量の計算を必要とする場所
SC3による並列処理

- Active領域1つだけだとSC3の並列性不足のため複数領域をまとめて処理
- 一番上の単位として染色体単位での並列化も行う

ZettaVEGAの性能

Precision FDA challenge V1 HG001のFASTQとGRCh38 リファレンスを用いて、HaplotypeCallerの出力をgVCFモードにして解析を行った (100Gbpのデータは通常vcfモードで実行)。

Pipeline	Input Data	Input Data size (Gbp)	BWA-MEM time (min)	Post-processing time (min)	HaplotypeCaller time (min)	Total time (min)
GATK Best Practice (CPU)	Precision FDA challenge V1 HG001	160.70	127.49	1,175.83	1,596.60	2,899.92
ZettaVEGA	Precision FDA challenge V1 HG001	160.70	23.62	2.61	16.89	43.12
ZettaVEGA	Precision FDA challenge V1 HG001	100.00	14.75	1.79	7.96 (vcf mode)	24.24

Input Data	Original only	ZettaVEGA only	Called by both	Match rate(%)
Precision FDA challenge V1 HG001	136	142	148,966,045	99.9998

BWA-MEMでは、PEZY-SC3への移植とリード単位での超並列処理、後処理に関では、ZettaScalerの最大2TBのメインメモリを活かした中間ファイルのオンメモリ化、HaplotypeCallerでは、PairHMMやSWAのPEZY-SC3への計算のオフロード、これらを行うことで処理時間の 48 hr (2,899 min) -> 43 min への大幅な短縮が可能となった。

ZettaVEGAの性能

Hardware	Software	解析時間 (分)
ZettaScaler 3.0 (PEZY-SC3 x 4)	ZettaVEGA v2.20.3	32.35
高火力 PHY(H100x8)	Parabricks v4.2.1	37.87
遺伝研 igt (V100x4)	Parabricks v4.2.1	90.05

Parabricksの速度のデータは遺伝研スパコンのサイトのベンチマーク結果より引用
https://sc.ddbj.nig.ac.jp/advanced_guides/benchmark/benchmark_parabricks
このベンチマークで使用されているNCGM-WGSpipelineパイプラインを実装し、
入力として 1000 人ゲノムプロジェクトから 20 サンプル (NA18941、NA18945、NA18946、NA18952、
NA18953、NA18957、NA18960、NA18964、NA18969、NA18971、NA18972、NA18976、
NA18983、NA18988、NA18990、NA18991、NA18995、NA19001、NA19002、NA19006) を選択し
GRCh38リファレンスを用いて同条件で解析した。

Contents

- メニーコアプロセッサPEZY-SC3とZettaScaler3.0サーバーについて
- 高速ヒトゲノム解析システムZettaVEGAについて
- **ZettaVEGAによる大規模ヒトゲノム解析**
- PEZY-SCシリーズのロードマップ

大規模ヒトゲノム解析と解析データの公開

国立遺伝学研究所 生命情報・DDBJセンター、株式会社ゲノムアナリティクスジャパンと合同で、2022年に公開された高精度なヒトゲノムT2T-CHM13をリファレンスゲノムとして3,193人分の大規模ゲノム解析を実施しました。解析結果についてはすべて公開しています。合計2基のZettaScaler3.0を使用し、平均の処理時間は一人あたり42.09分でした。

コンソーシアム	サンプル数	合計計算時間 (一人あたり平均 [分])	合計データ量 [TB] (一人あたり平均 [GB])
1000 人ゲノムプロジェクト	2,504	1,651 時間 (39.56 分)	123 TB (49 GB)
ヒトゲノム多様性プロジェクト	689	589 時間 (51.25 分)	54 TB (78 GB)
合計	3,193	2,240 時間 (42.09 分)	177 TB (55 GB)

プレスリリース:国産コンピュータ・プロセッサで実現したヒトゲノム解析の革新的な高速化と制約無しの完全公開

https://www.nig.ac.jp/nig/ja/2024/12/information_ja/pr20241226.html

データ公開サイト: <https://ddbj.nig.ac.jp/public/public-human-genomes/CHM13/>

今年稼働予定の次期遺伝研スーパーコンピューターに、ZettaScaler3.0 が2ノード導入される予定です

<https://sc.ddbj.nig.ac.jp/guides/hardware/hardware2025/>

Contents

- メニーコアプロセッサPEZY-SC3とZettaScaler3.0サーバーについて
- 高速ヒトゲノム解析システムZettaVEGAについて
- ZettaVEGAによる大規模ヒトゲノム解析
- PEZY-SCシリーズのロードマップ

PEZY-SCシリーズのロードマップ

	PEZY-SC	PEZY-SC2	PEZY-SC3	PEZY-SC3S	PEZY-SC4S	PEZY-SC5	
Release	Aug-14	Apr-17	Apr-21	Apr-22	2025 Q4	(2027)	
Process	28nm	16nm	7nm	7nm	5nm	3nm or finer	
Die Size	412m2	620m2	786m2	109m2	556m2	800m2	
Compute Cores(PE)	1,024	2,048	4,096	512	2,048	16,384	
Core Voltage	0.9V	0.8V	0.75V	0.75V	0.75V	0.75V	
Frequency	733MHz	1GHz	1.2GHz	1.0GHz	1.5GHz	2.0GHz	
Performance	64bit	0.75TFLOPS	4.1 TFLOPS	19.66 TFLOPS	2TFLOPS	24.5 TFLOPS	65.5TFLOPS
	32bit	1.5TFLOPS	8.2 TFLOPS	39.32 TFLOPS	4TFLOPS	49.0 TFLOPS	131TFLOPS
	16bit	3.0TFLOPS	16.4 TFLOPS	78.64 TFLOPS	8TFLOPS	589 TFLOPS(BF16)	4,000TFLOPS(BF16)
Cache Memory	L1	2KB	2kB	4KB	4KB	4KB	4KB
	L2	64KB	64KB	64KB	64KB	64KB	64KB
	LLC	16MB	40MB	64MB	-	64MB	64MB
External Memory		DDR3/4	DDR4	HBM2	HBM2	HBM3	HBM4
	Capacity	32GB	64GB	32GB	8GB	64GB/96GB/144GB	216GB
	Bandwidth	153.6GB/s	104GB/s	1.23TB/sec	0.6TB/s	3.2TB/sec	19.2TB/s
Host I/F		PCIe Gen3(x16)	PCIe Gen3(x16)	PCIe Gen4(x16)	PCIe Gen4(x4)	PCIe Gen5(x16)	PCIe Gen7(x32)
	Bandwidth	16GB/s	16GB/s	32GB/s	8GB/s	64GB/s	512GB/s
Power Consumption	100W	200W	470W	80W	270W	-	

ZettaScaler-4.0

	ZettaScaler-3.0 Server Unit	ZettaScaler-4.0 Server Unit (Temporary specifications)
Release	21-Apr	2025 Q4
Baseboard	EPX-BASE2 x 1	EPX-BASE3 x 1
CPU	EPYC 7713P 64Core 2.0GHz x 1	EPYC 9555P 64Core 3.1GHz x 1
Accelerator	PEZY Computing MOD-SC3H (PCIe Gen4x16 bus) x 4	PEZY Computing MOD-SC4H (PCIe Gen5x16 bus) x 4
Main memory	DDR4 ECC Registered 3200MHz SDRAM 64GB x 16	DDR5 ECC Registered 4800MHz SDRAM 96GB x 24
Storage	M.2 NVMe SSD 2TB x 4	M.2 NVMe SSD 2TB x 4
OS	AlmaLinux 9 or Ubuntu24.04LTS	AlmaLinux 9 or Ubuntu24.04LTS

ZettaScaler-4.0におけるヒト全ゲノム解析の処理速度は
ZettaScaler-3.0に比較して**約2倍高速化**される見込みです。

BF16での演算性能が大幅に上昇するため、DeepVariantやDeepSomaticなどを含めた深層学習関連のアプリケーションや学習が高速化されます。
ZettaVEGAにおいては、がんゲノム解析で使われるパイプラインで使用されるソフトウェアの移植などを行い機能強化を進めます。

PEZY-SCシリーズは汎用的なアクセラレータであり、ゲノム解析以外の様々な演算の高速化処理が可能です。

謝辞

ZettaVEGAの開発

OSS developers of BWA, samtools, biobambam2, Picard and GATK

大規模ヒトゲノム解析

国立遺伝学研究所

小笠原 理 先生

有田 正規 先生

株式会社ゲノムアナリティクスジャパン

八谷 剛史 様

野川 駿 様

Some images are from TogoTV

(© 2016 DBCLS TogoTV, CC-BY-4.0 <https://creativecommons.org/licenses/by/4.0/deed.ja>).