

Sentieon Genomics Tools



Sentieon Genomics Toolsは、次世代シーケンサーを用いた遺伝子変異解析における、ゲノムへのマッピングや変異検出を行うソフトウェアです。次世代シーケンサーにおけるバイオインフォマティクス解析のゴールドスタンダードである、BWA-GATKおよびMuTect、MuTect2と同じ数学モデルの解析パイプラインを、より高速に実行することが可能です。

特長

Sentieon Genomics Toolsでは、Broad Instituteが開発した遺伝子変異検出用のバイオインフォマティクス解析パイプラインBWA-GATK Best Practiceと同じ数学モデルを使用した「Sentieon DNaseq」と、腫瘍/正常のペアサンプルデータを使用した、体細胞変異検出用のパイプラインMuTect、MuTect2に対応した「Sentieon TNseq」の2種類の解析パイプラインが使用可能です。どちらのパイプラインも、一般的なCPUベースのシステムで動作し、またFASTQファイルからVCFファイル作成までに必要とする計算時間の、10倍以上の高速化を実現しています。また、コマンドライン型インターフェースを採用しており、複数サンプルデータの一括処理や、ユーザー独自の解析パイプラインの構築に適しています。

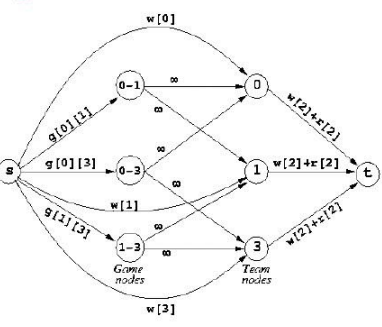


$$\frac{\partial}{\partial \theta} \left(\frac{1}{2} \sum_{i=1}^n \sum_{j=1}^n \frac{1}{\sigma^2} (y_{ij} - \mu_{ij})^2 \right) = -\sum_{i=1}^n \sum_{j=1}^n \frac{1}{\sigma^2} (y_{ij} - \mu_{ij})$$

$$= -\sum_{i=1}^n \sum_{j=1}^n \frac{1}{\sigma^2} \left(y_{ij} - \sum_{k=1}^K \beta_k x_{ijk} \right)$$

$$= -\sum_{i=1}^n \sum_{j=1}^n \frac{1}{\sigma^2} \left(y_{ij} - \sum_{k=1}^K \beta_k \left(\sum_{l=1}^L x_{ijl} \right) \right)$$

$$= -\sum_{i=1}^n \sum_{j=1}^n \frac{1}{\sigma^2} \left(y_{ij} - \sum_{k=1}^K \beta_k \left(\sum_{l=1}^L x_{ijl} \right) \right)$$



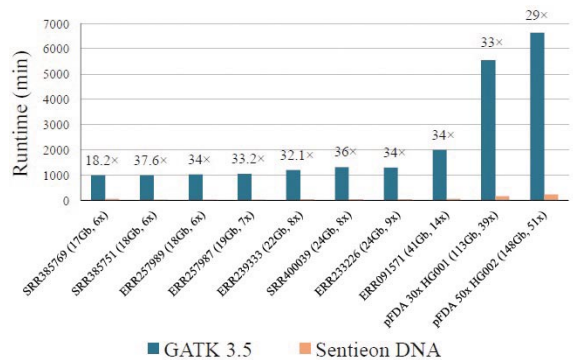
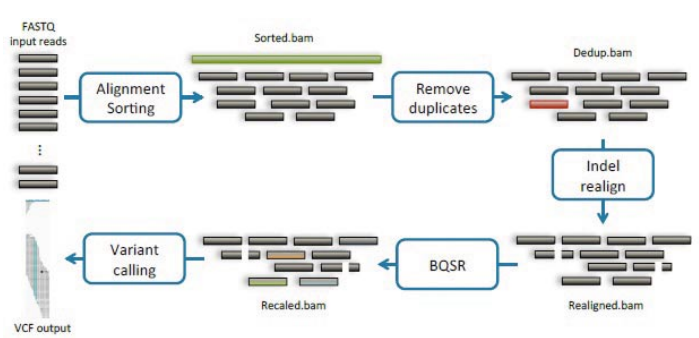
<http://www.cs.princeton.edu/courses/archive/spr05/cos226/assignments/baseball/>

```
1 /* This line basically imports the "stdio" header file, part of
2  * the standard library. It provides input and output functional
3  * to the program.
4  */
5 #include <stdio.h>
6
7 /*
8  * Function (method) declaration. This outputs "Hello, world" to
9  * standard output when invoked.
10 */
11 void sayHello() {
12     // printf() in C outputs the specified text (with optional
13     // formatting options) when invoked.
14     printf("Hello, world!");
15 }
16
17 /*
18  * This is a "main function". The compiled program will run the
19  * defined here.
20 */
21 void main() {
22     // Invoke the sayHello() function.
23     sayHello();
24 }
```

http://www.wikiwand.com/en/Programming_language

● Sentieon DNaseq

Sentieon DNaseqは、Broad InstituteのBWA-GATK HaplotypeCaller 3.x Best Practiceのバイオインフォマティクスパイプラインをもとに開発されており、同パイプラインと同じ数学モデルを用いて、リファレンスゲノム配列へのリード配列のマッピングから、変異の検出までを行うことが可能です。ただし計算アルゴリズムの効率化が行われており、FASTQファイルからVCFファイル作成の計算処理の場合は10倍、BAMファイルからVCFファイル作成の場合は20~50倍の高速化を実現しています。また100,000ものサンプルデータの場合でも、中間ファイルの統合を行わずに、効率的な計算が可能です。



*Server specs: 32 core 2.4 GHz Intel Xenon server, 64 GB memory

デモ版ダウンロードをご希望の場合は弊社までご連絡ください

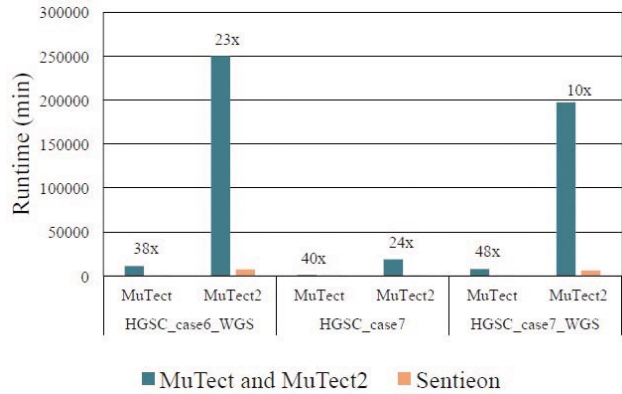
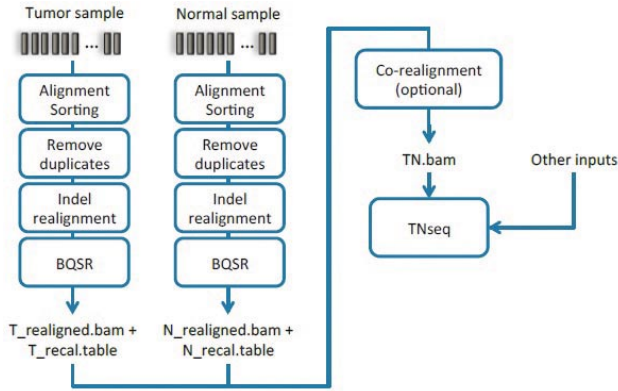
ソニー

● Sentieon Genomics Tools



● Sentieon TNseq

Sentieon TNseqは、Broad Instituteの体細胞変異検出用パイプラインMuTect (SNV), MuTect2 (SNVとINDEL)と同じ数学モデルを使用した解析パイプラインです。腫瘍/正常の各サンプルのシーケンスデータを用いて、効率化された計算アルゴリズムで、MuTectおよびMuTect2の10倍以上の計算速度で、体細胞変異の抽出を効率的に行います。



● 受賞成績

Sentieon Genomics Toolsは、「ICGC-TCGA Dream Mutation Calling Challenge」「precisionFDA Consistency Challenge」「precisionFDA Truth Challenge」の3つのコンテストにおいて、解析結果データの一致率や正確性の評価で、トップの成績を収めています。

precisionFDA Consistency Challenge Reproducibility
F1-score(%) between runs and between samples

| | Garvan vs. Garvan rerun | | | Garvan vs HLI | | |
|-------------------------|-------------------------|------|----------|---------------|--------|----------|
| | FP | FN | F1-score | FP | FN | F1-score |
| Sentieon by UNM* | 0 | 0 | 100 | 134633 | 315831 | 95.07 |
| Sentieon (dual mapping) | 0 | 0 | 100 | 107302 | 295568 | 95.53 |
| ISAAC** | 0 | 0 | 100 | 112255 | 266952 | 95.32 |
| Genallice** | 3621 | 3673 | 99.92 | 286306 | 433504 | 92.08 |
| Edico Dragen** | 3147 | 3216 | 99.93 | 161213 | 315611 | 94.87 |

precisionFDA Consistency Challenge Accuracy
F1-score(%) to NIST truth set

| | Garvan sample | | | HLI sample | | |
|-------------------------|---------------|----------|------------|------------|----------|------------|
| | All | SNP only | Indel only | All | SNP only | Indel only |
| Sentieon by UNM* | 99.39 | 99.86 | 95.85 | 98.97 | 99.73 | 92.94 |
| Sentieon (dual mapping) | 99.47 | 99.88 | 96.37 | 99.06 | 99.77 | 93.46 |
| ISAAC** | 97.29 | 98.57 | 86.66 | 96.34 | 97.99 | 82.10 |
| Genallice** | 98.04 | 99.20 | 89.10 | 97.25 | 98.83 | 84.66 |
| Edico Dragen** | 99.25 | 99.74 | 95.49 | 98.85 | 99.62 | 92.74 |

* Sentieon standard pipeline results from Jeremy Edwards (University of New Mexico) submission

** Edico, Genallice (MAP 2.2.0) and Isaac (aligner v01.14.07.14 & variant caller v2.0.13) run results from Changhoon Kim's (Macrogen Clinical Laboratory) submission

ライセンスタイプ

Sentieon Genomics Toolsでは、解析を行うサンプル数に合わせて、3種類のライセンスタイプがあります。下記の表は、ライセンスごとの1年間に解析を実行できるおよそのサンプル数です。

| | Whole Genome Sequencing | Whole Exome Sequencing | Gene Panel Sequencing |
|--------|-------------------------|------------------------|-----------------------|
| Tier 1 | 120 | 1,000 | 24,000 |
| Tier 2 | 600 | 5,000 | 120,000 |
| Tier 3 | 2,400 | 20,000 | Not appropriate |

System Requirements

- 16 GB+ of RAM (32 GB for Servers)
- 8+ CPU Cores
- 64-bit Windows 7 or later, Mac OS X 10.9 or later, Linux Ubuntu 14.04 or later (64-bit only), RHEL 6 or later (64-bit only), or equivalently CentOS 6 or later (64-bit only)

Price

| 製品名 | ライセンスタイプ | 初回購入税別価格 | カタログ# | 初回購入税別価格 | カタログ# |
|-------------------------|----------|----------|----------|----------|-----------|
| Sentieon Genomics Tools | Tier 1 | お問い合わせ | ST-T1-AF | お問い合わせ | ST-T1-AFM |
| | Tier 2 | お問い合わせ | ST-T2-AF | お問い合わせ | ST-T2-AFM |
| | Tier 3 | お問い合わせ | ST-T3-AF | お問い合わせ | ST-T3-AFM |

* ソフトウェアは年間ライセンスとなります。2年目以降の使用には、別途年間更新費用が必要です。

* テクニカルサポートは、電話、あるいは、電子メールでの対応のみとなります。