

Illumina Connected Analytics

バイオインフォマティクス
オペレーションを変革

- シーケンスシステムとの効率的な統合
- サンプル処理のニーズを満たすスケーラブルなスループット
- ワークフローを選択、構築、カスタマイズするための直感的なインターフェース

はじめに

次世代シーケンサー (NGS) テクノロジーの進歩により、ライフサイエンス研究や臨床研究の実施速度は劇的に変化してきました。シーケンスの高速化とコスト低下により、データ生成能力はデータから生物学的洞察と臨床的洞察を抽出する能力をはるかに上回ると考えられます。パイプラインの開発と実装、インフォマティクスワークフローの拡張、安全なデータ管理の維持といった課題に対処するには、柔軟で包括的なプラットフォームが必要です。Illumina Connected Analyticsでは、ユーザーがデータプライバシー、セキュリティおよびコンプライアンスを維持しながら、柔軟な解析パイプラインを構築、バージョンアップ、展開することができます。

Illumina Connected Analyticsは、インフォマティクスの運用と科学的洞察を推し進める、セキュアなクラウドベースのバイオインフォマティクスデータプラットフォームです。(図1、表1)。Illumina Connected Analyticsを使用すると、次のようなワークフローを作成できます。

- イルミナのシステムおよびソフトウェアと接続できるため、効率的な統合が可能
- スケーラブルであるため、膨大なデータセットを管理、解析、探索する機能がある
- 柔軟性が高く、DRAGEN™アプリケーションとカスタム解析パイプラインを使用してワークフローを構築およびカスタマイズ可能

効率化されたワークフロー

Illumina Connected Analyticsはイルミナのシーケンスシステムを用いてNGS研究を実施しているラボにとって、中心的なコンポーネントです。クラウドコンピューティングによるリソースの順応性を生かし、複雑なシングルセルプロジェクトの何万もの細胞の不定期なスクリーニングから、集団規模の全ゲノムシーケンスまで、同一のアーキテクチャーを用いて、あらゆる規模でオペレーションをサポートします。ユーザーはシームレスに手持ちの機器をIllumina Connected Analyticsと統合できます。

Illumina Connected Analytics内部では、指定したワークフローに応じて、すぐに使えるDRAGEN二次解析パイプラインまたはカスタムパイプラインでデータを自動的に解析できます。幅広い解析オプションは、品質管理からデータアグリゲーションや高度なデータサイエンスツールにまで及び、迅速かつ拡張性のあるデータ処理を実現します。Illumina Connected Analyticsは、豊富なRESTfulアプリケーションプログラムインターフェース (API) およびコマンドラインインターフェース (CLI) ツールを備え、拡張性のあるプラットフォームを提供します。これらのAPIは、データのライフサイクル全体で転送、アクセス、使用されるときにワークフロー効率を最大にします。

表1: Illumina Connected Analyticsの概要

	特長	利点
セキュリティおよびプライバシー	コンプライアンス	ローカル、地域、GDPR ² 、HIPAA ³ 、ISO 27001/27701 ^{4,5} およびISO 13485 ⁶ を含む世界的な規制基準に遵守
	セキュリティ管理	厳格なデータ分離、「転送中」(TLS 1.2) および「保管中」(AES 256) のデータ暗号化を維持
	監査証跡	アクセス監視のためのアクティビティとイベントログの追跡を維持
	多要素認証 (MFA)	システムアクセスを管理および制御するために施設承認情報を適用
リソース	要望に応じて規模を拡大	現在要求されるレベルのクラウドストレージと計算ニーズに応じた規模拡大
	要望に応じた計算リソース	パイプラインエンジン中の計算リソースのみの支払いで、費用を削減
管理	プロジェクトおよびユーザーの管理	細やかなプライバシーに対するユーザーアクセスとアクティビティの管理
	データ共有	データサイロを橋渡ししてグローバルなコラボレーションを促進し、データ配信をサポート
	データアーカイブ	アーカイブストレージ階層による長期データ管理のコスト削減
	プライベートAWSクラウドバケットの持ち込み	プライベートに管理されたAWSクラウドアカウント内に保存されたデータにアクセス
ユーザービリティとアクセスシビリティ	シーケンスシステムの統合	イルミナのシーケンスシステムからIllumina Connected Analyticsのストレージと解析にデータをシームレスにストリーミング
	ツールおよびパイプライン	すぐに使えるDRAGENパイプラインを使う、または、既存のツールをインポート
	APIおよびCLI	CLIまたはRESTful APIを介してプログラマティックにプラットフォーム、ツール、データと通信
	直感的なグラフィカルインターフェース	グラフィカルWebインターフェースを介してデータとツールにアクセスし、RおよびPythonパッケージを使用してデータを表示
高度なツール	Docker、Nextflow およびCWLに対応	Common Workflow Languageでパイプラインを書き出し、クラウドで簡単に解析を起動
	JupyterLabとの統合	カスタムスクリプトを実行し、AI/機械学習モデルを構築およびトレーニングし、共同ノートブックワークスペースを通じてプラットフォーム内のデータと通信
	データアグリゲーションおよびクエリー	構造化されたマルチオミクスデータセットを整理して探索し、三次解析を強化

ILLUMINA CONNECTED ANALYTICS

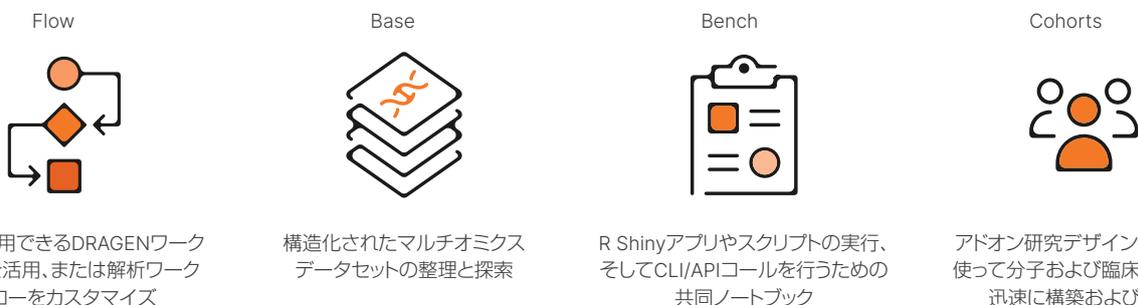


図1: Illumina Connected Analyticsモジュールコンポーネント: 安全でアクセス制御されたデータ管理環境内でデータを処理、集約、解析します。

BaseSpace™ Sequence Hub

BaseSpace Sequence HubはIllumina Connected Analyticsのすべてのサブスクリプションに含まれ、イルミナ装置から直接アクセスできます。BaseSpace Sequence Hubとの統合により、ランの設定、ラン品質の監視、シーケンスデータのクラウドへの直接ストリーミングが可能になり、ラボの効率が向上します。暗号化されたデータは装置からBaseSpace Sequence Hubに送られるため、安全な環境内で、キュレーションされた一連のアプリケーションを使って簡単にデータを管理し解析できます。

リードからデータへの変換

Illumina Connected Analyticsはデータ二次解析用のさまざまなオプションがあり、リードから結果取得までのワークフローを効率化します。Illumina Connected Analyticsは、既存のパイプラインや製品を使用し、カスタマイズされたパイプラインを構成するための柔軟性を備えているため、実質的にあらゆるインフォマティクスアプリケーションに対応できます。

すぐに使えるオプション

Illumina Connected Analyticsは、シーケンスデータの精確かつ効率的で包括的な二次解析を実施するDRAGEN二次解析¹へのアクセスをはじめとした、パワフルですぐに使えるデータ処理ツールおよびパイプラインを提供します (図2)。

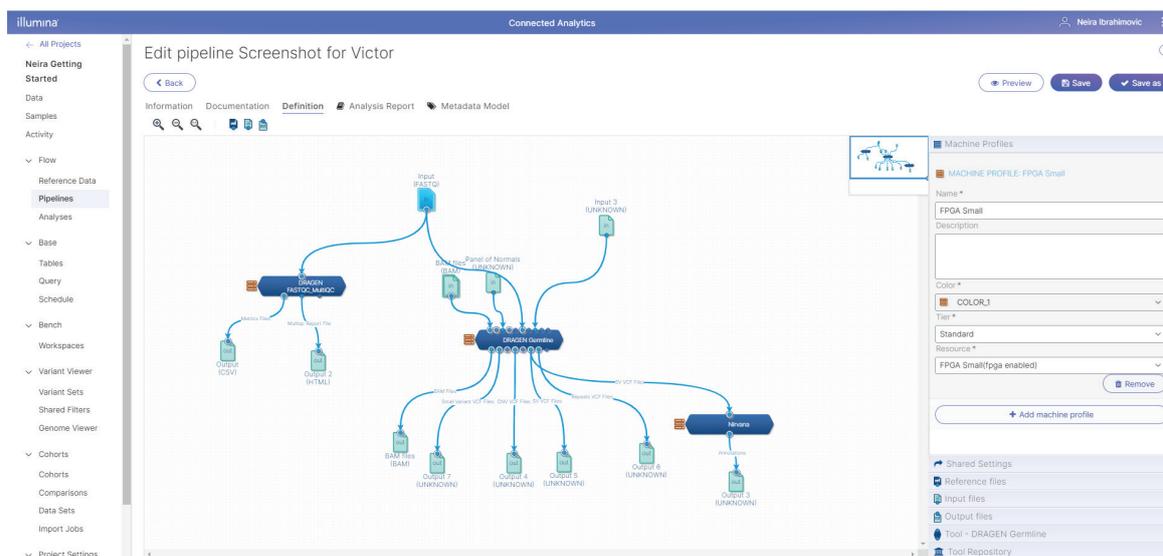


図2: Illumina Connected Analytics Flow: 解析パイプラインの柔軟な構築、展開、バージョン管理を行います。

パイプラインのカスタム化

バイオインフォマティクスはDocker Hubイメージリポジトリ⁷ から既存の解析ツールをインポートしたり、Nextflow、Common Workflow Language (CWL) およびグラフィカルパイプラインエディターを使って新しいパイプラインを構築したり編集したりできます。ラボオペレーターやその他のサイエンティストは、直感的にデザインされたユーザーインターフェースを使って簡単にパイプラインを起動することが可能です。

データ管理と制御

データの増加に伴い、個々のデータセットの価値を増大させるために、科学界の中でのデータ共有、再利用、統合に対応するためのインフラストラクチャーに対するニーズが高まっている時代が到来しています。このニーズに対処するために、Illumina Connected Analyticsはデータ管理におけるベストプラクティスの導入が実現するようにデザインされた、いくつかの性能を搭載しています。

アクセス制限

きめ細かいアクセス制限により、管理者が許可を設定し、既存の施設承認の利点を生かしてアクセスを制限することが可能です。監査ログはイベントおよび変更の記録として役立ち、プラットフォームを使用しながら、そのプラットフォームにアクセスしたときの各ユーザーのその行動を記録し、コンプライアンスと説明責任の実施を実現します。

オープンフォーマット

Illumina Connected Analyticsはデータに依存しないプラットフォームとして設計されています。分子データ、臨床データ、表現型データおよび画像などの非構造化データを含む多種類のデータ解析に対応します。

コラボレーション

Illumina Connected Analyticsはコンプライアンスを維持する方法で地理的境界線全体にわたるコラボレーションを後押しします。データおよびツールを、データ完全性とプライバシーを維持した形で瞬時に送信し、他のユーザーと共有できます。さらに、解析と共有のために、外部クラウドソースにホストされたデータツールと解析ツールをIllumina Connected Analyticsにインポートできます。

データアグリゲーションおよびクエリー

Illumina Connected Analyticsは複雑なアグリゲーションステップと統合ステップを自動化し、数百万個のサンプルのデータを網羅する機能的なナレッジマネジメントシステムを構築します。実質的にあらゆる種類のデータ、ジェノタイプ、表現型、メタデータ、アノテーション、その他の関連する情報を取得することができます。ユーザーは自身のデータモデルを定義し、自身のクエリーを書き出し、必要なデータセット間の関連を探索できます。Illumina Connected Analyticsに集積されたデータは、新規バイオマーカーの探索、患者集団の層別化、経時的なアッセイ性能のモニタリングなどに利用できる豊富な情報を含んでいます。

Cohortsを使用して臨床ゲノムデータを大規模に探索

Illumina Connected Analytics Cohortsは、コホートを迅速に構築および探索するための試験デザインツールです (図3)。Illumina Connected Analytics Cohortsは、ゲノムの発見から意義のある研究までの道のりを加速し、マルチオミクスおよび表現型データを1カ所に集約することで、研究設計に関するより適切な意思決定を可能にします。

The Cancer Genome Atlas (TCGA)、Broad Rare Genomes Project、1000人ゲノムプロジェクト、Gene Expression Omnibusなどの調和された公開データセットを組み込んでコホートを拡張します。統合されたコホートブラウザを活用することで、数週間ではなく数分で仮説を構築できます。正確性の高いキュレーションされた公開データセットシンプルな視覚ツールを使用してコホートを選択し、視覚化することで、複雑な探索を実行する必要がなくなります。

洞察を推し進めるためのセキュアなノートブック環境

アルゴリズムの開発とカスタム化は、詳細なデータ探索をサポートするためのIllumina Connected Analyticsの重要なコンポーネントです。データサイエンティストは、定評のあるJupyterノートブック、PythonおよびRを活用した、インタラクティブなプログラミングモジュールにより、シームレスかつセキュアな環境で集積されたデータを解析できるようになります (図4)。

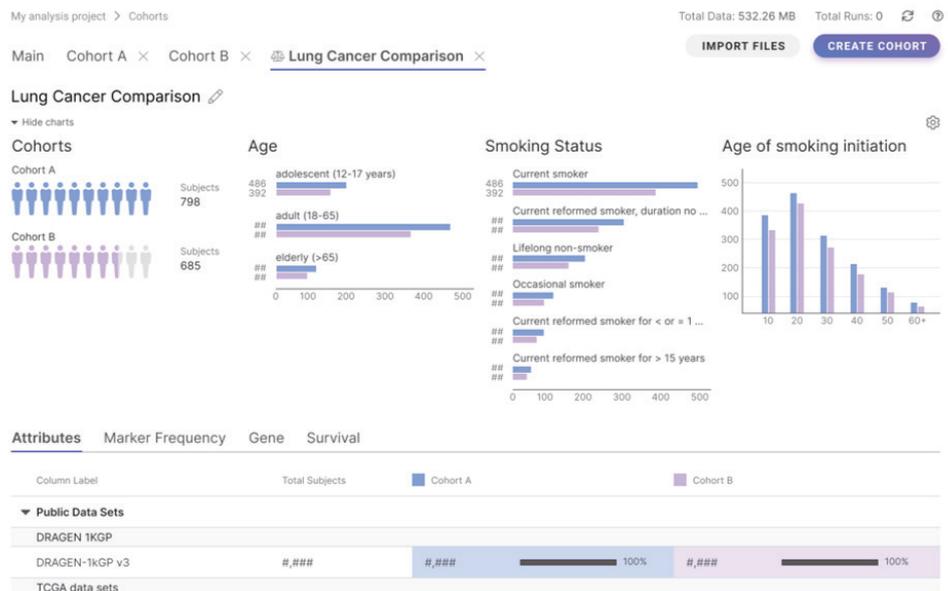


図3: Illumina Connected Analytics Cohorts: Illumina Connected Analyticsのこのアドオンモジュールを使用して、分子データと臨床データを迅速に構築し調査します。

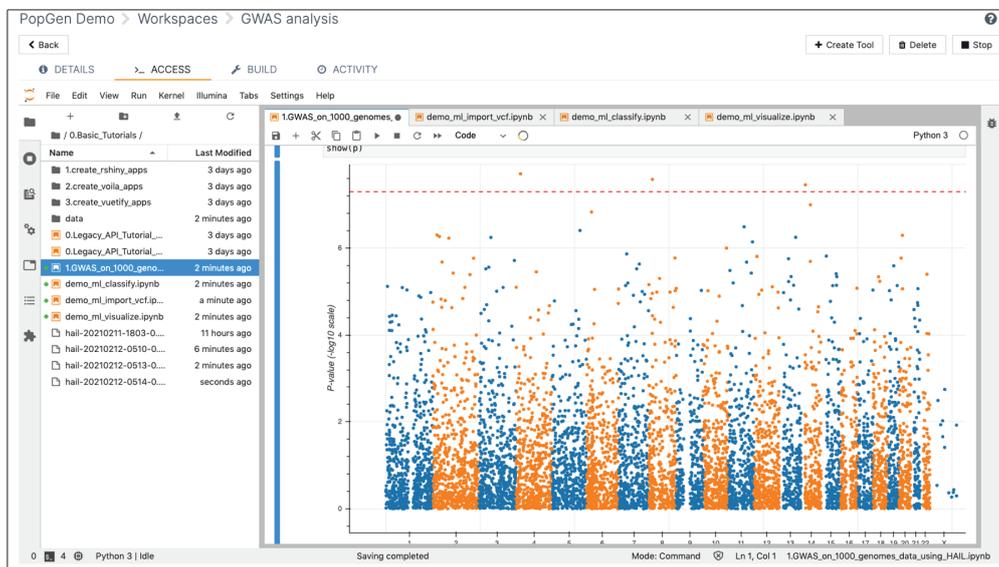


図4: Illumina Connected Analytics Bench: Jupyterノートブック、Python、Rなどの柔軟でインタラクティブなコンピューティングツールがIllumina Connected Analytics Benchに組み込まれています。

メソッドおよびアルゴリズムの開発段階では、サンドボックス環境でパイプラインを開発、変更できます。ユーザーは標準的なライブラリにアクセスし、TensorFlow⁸ またはscikit-learn⁹ などの自身のカスタムライブラリにそれらのライブラリを簡単に持ち込むことができ、Illumina Connected Analytics内で共同作業によってデータを処理するためのカスタムかつ複雑なスクリプトを構築できます。生産段階に移行する準備ができたときに、Illumina Connected Analyticsによってノートブックをツールに変換できます。この後、これらのツールはIllumina Connected Analyticsツールのリポジトリで使用できるようになり、生産パイプラインに組み込まれることになります。

核となるセキュリティとコンプライアンス

研究、臨床治療、人間の診断のためにゲノムクスデータを操作する場合、セキュリティは最も重要です。Illumina Connected Analyticsは次のようなさまざまなデジタル対策および管理手段を使用し、最も厳しいデータセキュリティ要件をも満たします。

- シーケンス機器からアップロードされたデータをAES256標準¹⁰ を使って暗号化し、トランスファーレイヤーセキュリティ (TLS) で保護します。
- Illumina Connected Analytics内のデータはアマゾンウェブサービス (AWS) 上でホストされ、AWS Well-Architectedのベストプラクティスを用いてさまざまな業界認証セキュリティ基準への遵守を維持します。¹¹
- Illumina Connected Analyticsは、特定の地理的領域内でゲノムデータを維持するための現地の規制に準拠できるように、世界中の10を超えるAWSリージョンでホストされています。
- この認証サービスはセキュリティアサーションマークアップ言語 (SAML) 2.0でサポートされ、施設のユーザーとパスワードを管理します (オプション)。
- 監査レポートは、データ由来のトレーサビリティをサポートします。

Illumina Connected Analyticsは、規制された環境でオペレーションを行う、次のような厳しい環境を遵守する必要があるカスタマーもサポートします。Illumina Connected Analyticsは、イルミナの品質管理システム (QMS) に基づくイルミナソフトウェア開発ライフサ

イクルプロセスに従って開発されました。また、イルミナQMS内のプロセスでは、次のような業界のベストプラクティスと関連スタンダードが採用されています。

- 国際標準化機構 (ISO) ISO 27001:2013情報セキュリティ管理システム⁴ およびISO 27701:2019セキュリティ技術⁵
- ISO 13485組織品質管理システム (QMS) のベストプラクティス⁶
- 一般データ保護規則 (GDPR)²
- 医療保険の携行性と責任に関する法律 (HIPAA)³
- 地域の規制およびコンプライアンス要件に対応する保証されたデータレジデンシー

製品情報

製品	カタログ番号
ICA Enterprise Annual Subscription	20038994
ICA Professional Annual Subscription	20044876
ICA Training and Onboarding	20049422
ICA Cohorts Annual Subscription	20065842
Illumina Analytics - 1 iCredit	20042038
ICA Enterprise Srvc and Compliance Add-on	20066830
略語: ICA、Illumina Connected Analytics	

詳細はこちら

[Illumina Connected Analytics](#)

参考文献

1. Illumina. Illumina DRAGEN secondary analysis. www.illumina.com/products/by-type/informatics-products/dragen-secondary-analysis.html. Accessed March 1, 2024.
2. General Data Protection Regulation (GDPR) Compliance Guidelines. GDPR website. gdpr.eu. Accessed March 1, 2024.
3. US Department of Health and Human Services. Health Information Privacy. HHS website. hhs.gov/hipaa/index.html. Accessed March 1, 2024.
4. International Organization for Standardization. ISO-ISO/IEC 27001—Information security management. ISO website. iso.org/isoiec-27001-information-security.html. Accessed March 1, 2024.
5. International Organization for Standardization. ISO/IEC 27701:2019—Security techniques. iso.org/standard/71670.html. Accessed March 1, 2024.
6. International Organization for Standardization. ISO 13485—Medical devices. ISO website. iso.org/iso-13485-medical-devices.html. Accessed March 26, 2024.
7. Docker. Docker Hub Image Container Library. <https://hub.docker.com>. Accessed March 26, 2024.
8. TensorFlow. TensorFlow website. tensorflow.org. Accessed March 1, 2024.
9. scikit-learn: machine learning in Python. scikit-learn website. scikit-learn.org/stable/. Accessed March 1, 2024.
10. National Institute of Standards and Technology. Advanced Encryption Standard (AES). nist.gov/publications/advanced-encryption-standard-aes-0. Accessed March 26, 2024.
11. Cloud Security—Amazon Web Services (AWS). Amazon website. <https://aws.amazon.com/security/>. Accessed March 1, 2024.
12. Illumina. iCredits for Data Storage & Analysis. illumina.com/products/by-type/informatics-products/icredits.html. Accessed March 1, 2024.

イルミナ株式会社

〒108-0014 東京都港区芝 5-36-7 三田ベルジュビル 22階
Tel (03) 4578-2800 Fax (03) 4578-2810
jp.illumina.com

 www.facebook.com/illuminakk

販売店

本製品の使用目的は研究に限定されます。診断での使用はできません。 販売条件：jp.illumina.com/tc

© 2024 Illumina, Inc. All rights reserved.
すべての商標および登録商標は、Illumina, Inc. または各所有者に帰属します。
商標および登録商標の詳細は jp.illumina.com/company/legal.html をご覧ください。
予告なしに仕様および希望販売価格を変更する場合があります。