

希少疾患の解析に情報をもたらし、バリエーションの解釈を強化

TruSight™ Software Suite のユースケース

はじめに

次世代シーケンス（NGS）テクノロジーを使用した全ゲノムシーケンス（WGS）は、遺伝性疾患に関連するバリエーションを調査するための強力なツールです。ゲノム全体を高解像度で偏りなく見渡し、遺伝性疾患の原因となるバリエーションを発見できます。ただし、WGSによって生成される膨大な量のデータには、生のシーケンスデータを意味のある解釈可能な結果に効率的に変換できる包括的なデータ解析ツールが必要です。こうした課題に対処するために、イルミナでは TruSight Software Suite を提供しています。

この仮定のユースケースでは、ラボアナリストが TruSight Software Suite を使用して遺伝性疾患に関連するバリエーションを特定するワークフローを説明します（図 1）。

バリエーション解析のケース受領とリクエスト

ラボアナリストは WGS バリエーション解析の新しい依頼を受け取ります。このケースは、男性の発端者と影響を受けていない両親で構成されるトリオデータセットの例（家族 1）です（表 1）。発端者および各親の DNA サンプルからライブラリーが調製され、NovaSeq™ 6000 システムでシーケンスされます。その後、アナリストは TruSight Software Suite で作業を開始し、発端者の表現型の特徴に対応する疾患関連のバリエーションを見つけます。

表 1：データセットのケースの説明の例

発端者				
ケース	性別	表現型	母親	父親
家族 1	男性	- 夜盲症 - 視神経乳頭蒼白	影響を受けていない	影響を受けていない

症例管理

TruSight Software Suite の Case Management Portal を使用すると、ラボアナリストは、新しいケースの作成、データファイルのインポート、シーケンスデータファイルのケースへの関連付けを行えます。あるいは、これらの情報は、アプリケーションプログラミングインターフェース（API）を介して簡単にインポートすることもできます。次に、ラボアナリストは、家族構造、発端者の性別、発端者の表現型の特徴（オプション）、影響を受けている家族の状態など、ケースに関する家系情報（最大 5 人）を入力します（表 1）。ケースのステータスがリアルタイムで更新されて TruSight Software Suite に表示されるため、アナリストは解析ワークフロー全体で進行状況を監視できます。

シーケンスが完了すると、DRAGEN™ Bio-IT Platform によって、発端者および関連する家族のシーケンスアライメントとバリエーションコールが自動的に実行されます。包括的なバリエーションコールには、小さなバリエーション、コピー数バリエーション（CNV）、構造多型、ミトコンドリアバリエーション、リピート伸長などが含まれます。[Quality Control (QC) Metrics] タブで発端者と各家族の情報を確認した後、ラボアナリストはバリエーションのトリアージ、視覚化、解釈を進めます（図 1）。

バリエーショントリアージ

バリエーショントリアージを開始するには、[Interpretation] タブを開きます。バリエーションフィルタリングは、カスタム計画や事前に作成されたフィルター計画に従うことができます。TruSight Software Suite の家族ベースのフィルタリングを使用すると、発端者和其他の家族を比較できます。追加のフィルターオプションには、Genome Aggregation Database (gnomAD) などのソースからの集団内頻度、バリエーションの結果、サンプルパラメーター、

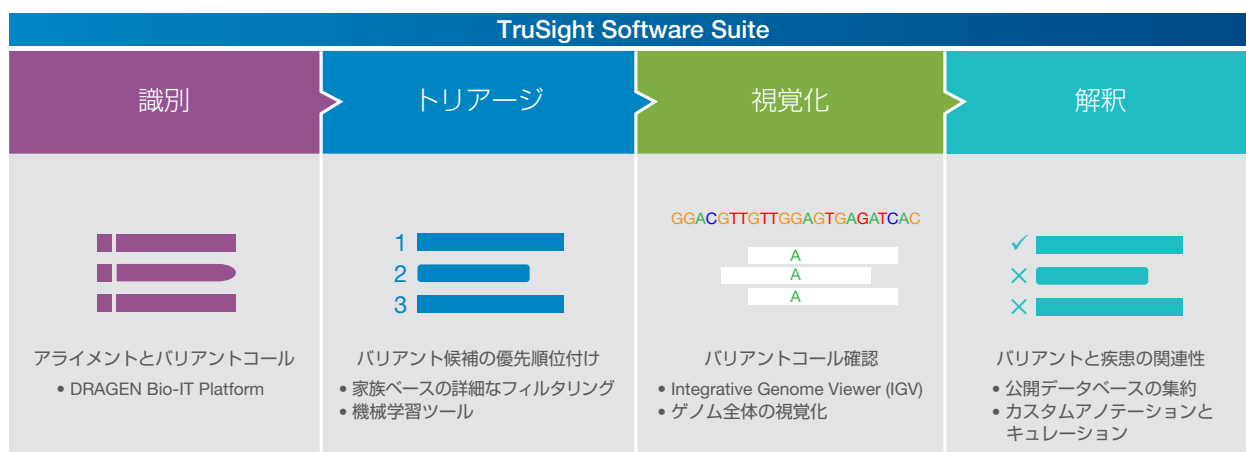


図 1：TruSight Software Suite でのバリエーション解析：DRAGEN プラットフォームがアライメントとバリエーションコールを実行した後、ユーザーはバリエーションのトリアージ、視覚化、および解釈を進めます。

TEST-CASE6

SUBJECT INFO INTERPRETATION REPORTS

Subject: PROBAND

BETA_PROTOCOL_FILTERS GENERAL DENOVO RECESSIVE SHORT NEGATIVE SECONDARY FINDINGS MOST LIKELY + NEW VIEW

Show Filters View variants by flag Displaying 28 of 5004596 Variants.

FLAG	STATUS	LINKED	IGV	CATEGORY	POSITION	GENE	OVERLAP	CONSEQUENCE	HGVS	ZYGOSITY	POP FREQ	LENGTH	CLINVAR
		CH		Small Variant	6 : 65622398	EYS	5	Stop Gained	c.2620C>T p.(Glu1076Val)	Heterozygous (Father)	0.000079 NFE gnomAD	1	P
		CH		Copy Number Variant	6 : 65985254-66085786	EYS (+1)	5	Transcript Truncation	Cop	Heterozygous (Mother)		100533	
				Run of Homozygosity	8 : 17453812-22542961	PDGFRL (+47)	5			Hemizygous		5089150	
				Run of Homozygosity	13 : 950631...-1017984...	DCT (+57)	5			Hemizygous		6735353	
		CH		Small Variant	10 : 126103...-126103...	OAT	4	Intron Variant	c.-29-2888...-29-2...	Heterozygous		14	
				Small Variant	20 : 5092253-5092254	TMEM230 (+1)	4	Splice Region Variant	Intr c.69-5-69-3dup...	Homozygous (Conflict)		3	
		CH		Small Variant	2 : 210640622	UNC80	1	Stop Gained	c.151C>T p.(Arg50Gln)	Heterozygous		1	P
		CH		Small Variant	10 : 14139140-14139141	FRMD4A	1	Intron Variant	c.45+232913_4...	Heterozygous		4	
		CH		Short Tandem Repeat	18 : 5325387-53253458	TCF4	1	Short Tandem Repeat Expansion		Heterozygous		57	
		CH		Small Variant	1 : 39879371	MACF1 (+1)		Intron Variant	c.9631-8675T>C	Heterozygous	0.000149 EAS gnomAD	1	
		CH		Small Variant	2 : 26022394	ASXL2		Missense Variant	c.263C>T p.(Pro88Leu)	Heterozygous	0.000027 NFE gnomAD	1	

図 2：バリエントのフィルタリングと優先順位付け：ラボアナリストは家族ベースのバリエントフィルタリングを行うことで、遺伝的バリエントと de novo バリエントを同定できます。図のように、バリエントグリッドには、バリエントのカテゴリ、染色体の位置、影響を受ける遺伝子、重複（バリエントの重複する表現型の数）、バリエントの結果、バリエントの集団内頻度（既知の場合）などに関する情報が表示されます。

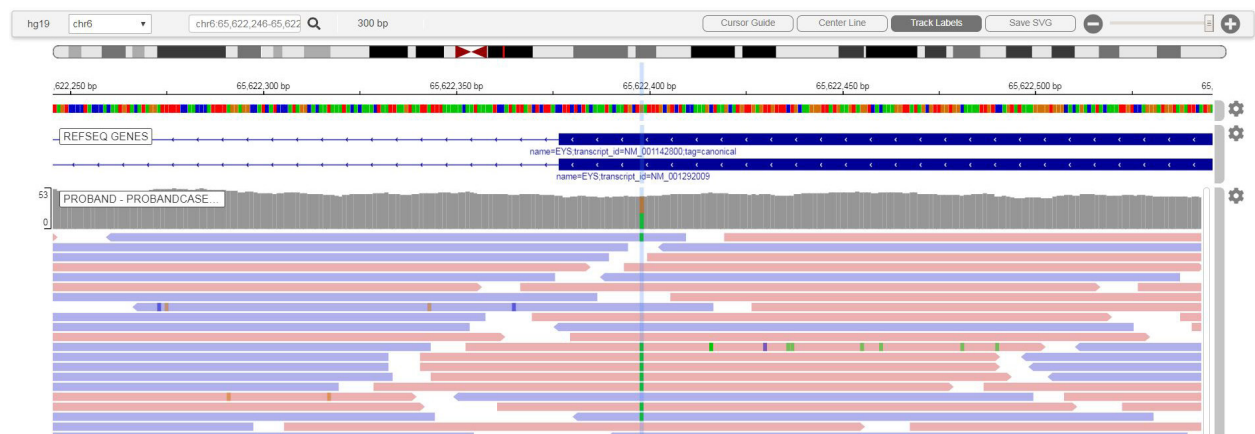


図 3：バリエントコール確認：ラボアナリストは、内蔵された IGV にアクセスして、詳細なバリエントの確認を行い、他のケースや外部データベースと比較できます。

Subject: PROBAND

BETA_PROTOCOL_FILTERS GENERAL DENOVO RECESSIVE SHORT NEGATIVE SECONDARY FINDINGS MOST LIKELY + NEW VIEW

Show Filters View variants by flag Displaying 28 of 5004596 Variants.

FLAG	STATUS	LINKED	IGV	CATEGORY	POSITION	GENE	OVERLAP	CONSEQUENCE	HGVS	ZYGOSITY	POP FREQ	LENGTH	CLINVAR
		CH		Small Variant	6 : 65622398	EYS	5	Stop Gained	c.2620C>T p.(Glu1076Val)	Heterozygous (Father)	0.000079 NFE gnomAD	1	P
		CH		Copy Number Variant	6 : 65985254-66085786	EYS (+1)	5	Transcript Truncation	Cop	Heterozygous (Mother)		100533	

Variant Details Transcript Gene Details Knowledge Network Send for Curation

SNV chr6:65622398

REF: G
ALT: A

Note
Add a note

Zygosity & Metrics

	PROBAND	MOTHER	FATHER
Affected	True	False	False
Genotype	0/1	0/0	0/1
VRF	0.475	0	0.4
Total Depth	40	33	45
Allele Depths	19	0	18
GT Quality	48	84	46
Filter	PASS	PASS	PASS
Quality Score	125	125	125

Population All Population

	GnomAD
ALL	0.00004
AFR	0.00000
AMR	0.00004
ASJ	0.00000
EAS	0.00000
EUR	0.00000
FIN	0.00000
NFE	0.00008
SAS	0.00000

SpliceAI (EYS)

Acceptor Gain	Donor Gain
Score: 0.00004	Score: 0.00000
Distance: 0.00000	Distance: 0.00000
Acceptor Loss	Donor Loss
Score: 0.1	Score: 0.21
Distance: 0.00000	Distance: 0.00000

図 4：バリエントの解釈とキュレーション：ユーザーはソフトウェア内で複数の外部データベースを検索し、バリエントの解釈に役立てることができます。

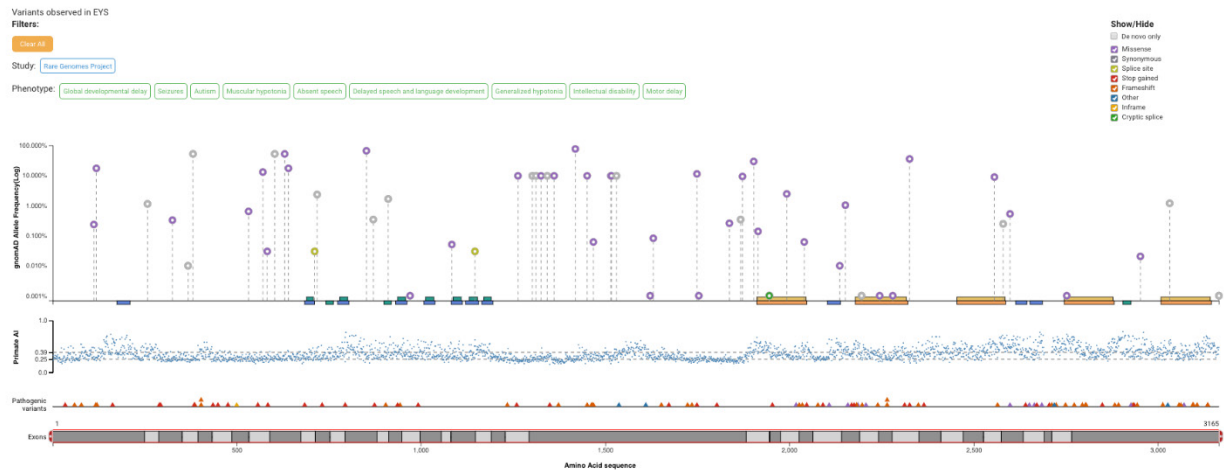


図 5 : CaseLog : ラボアナリストは、対象バリエントと対象遺伝子の両方で、集約データを視覚化できます。

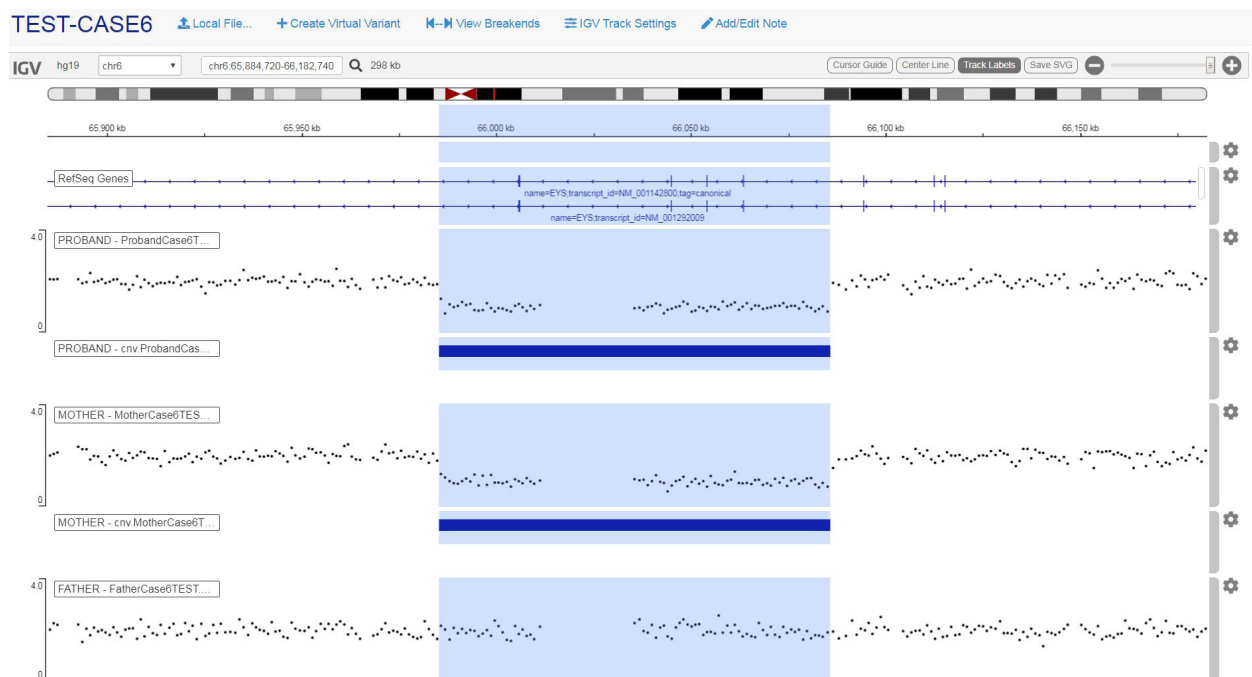


図 6 : 家族 1 の対象バリエント : ラボアナリストは TruSight Software Suite のツールを使用して、発端者の対象バリエント（小さなストップゲインバリエントと CNV で構成される EYS 遺伝子にある複合ヘテロ接合性バリエント）を同定しました。この画像は、ノーマライズされたカバレッジプロットと、発端者と母親の明確なコピー数減少（青色のボックスで強調表示）を示しています。

ClinVar 病源性などのフィルタリングが含まれます (図 2)。またアナリストは、SpliceAI や PrimateAI などの人工知能 (AI) ベースの機械学習を活用した、増え続けるツールを利用して、対象バリエーションをさらに細かく優先順位付けすることができます。これらのディープラーニングネットワークは、mRNA スプライスサイトとミスマッチバリエーションをそれぞれ偏りなく高精度で分類します^{1,2}。トリアージ後、ラボアナリストは数百万の無関係なバリエーションを除外し、視覚化と解釈では、そのケースで対象となる上位のバリエーションに焦点を当てることができます。

バリエーションの視覚化

アナリストは Integrative Genomics Viewer (IGV) などの組み込み視覚化ツールを使用して、ケース内の全被験者のリードレベルのアライメントやペアエンドリード、b アリル頻度、ノーマライズされたカバレッジトラックなどのゲノムデータを詳細に調べます (図 3)。バリエーションレベルの視覚化だけでなく、ズームアウトして染色体全体またはゲノム全体を表示し、大きな異常を確認することができます。バリエーションコールを確認したら、ケースの解釈を進めます。

バリエーションの解釈とキュレーション

最後に、現在のケースに関連する優先バリエーションを判断します。これを行うには、発端者のものに類似した重複する表現型と、遺伝子とバリエーションの関連性を調査します。TruSight Software は、Online Mendelian Inheritance in Man (ヒトにおけるオンラインメンデル遺伝学、OMIM) カタログや、ClinVar などの必要な外部データベースを [Variant Details] タブに集約・統合することで、これを可能にします (図 4)。集約されたデータを表示することで、バリエーションに関する貴重な洞察を得ることができ、別々のデータベースで何度もオンライン検索する必要がなくなります。また TruSight Software Suite には、American College of Medical Genetics (ACMG) が作成したバリエーション分類の標準的な用語が統合されているため、バリエーションの詳細や関係性の記録に役立ちます。これによりお客様は、遺伝子レベルの情報 (機能喪失バリエーションに対する耐性など) や、遺伝子と疾患の関係性の特徴に (必要な外部データベースを介して) アクセス・保存することができます。

また、[Note] フィールドなどの機能を使用して特定のバリエーションに関連するケース固有のメモを追加したり、[Comments] フィールドでバリエーションや遺伝子に関するケースに依存しない情報を記録したりすることもできるため、今後のケースで参照する際に役立つ場合があります。CaseLog は、各ケースの遺伝子、バリエーション、表現型情報を、非公開データセットと公開データセットの両方で表示・集約するために使用されます (図 5)。このユーザー固有のバリエーション情報データベースにより、科学界での新しい発見に基づいて過去のデータを再解析できます。

ケースの結果とレポート作成

解釈後、ラボアナリストは既知の疾患関連性を持つ発端者の対象バリエーションを特定し、キュレーションしました。その結果、複合ヘテロ接合性バリエーションが、小さなストップゲインバリエーションと、複数のエクソンの欠失をもたらす CNV を含む EYS 遺伝子で見つかりました (表 2、図 6)。EYS 遺伝子の変異は、網膜色素変性症 (RP) との関連性が知られています。RP は、暗視や周辺視野などの視力の不完全な喪失をもたらす遺伝性疾患です。

ラボアナリストは、新しいケースレポートを作成し、そのケースに関連する遺伝子とバリエーション関連性を追加します。レポートを送信すると、ソフトウェア内で追加のレビューと承認を行います。データ共有を容易にするために、レポートは PDF 形式または JSON 形式でダウンロードできます。これで Case Management Portal でのケースのステータスは、[Reported] に更新されます。

表 2: 家族 1 の発端者のバリエーション解釈

対象バリエーション	解釈
EYS の小さなバリエーション 6:65622398 (c.2620C>T; p.Gln874Ter) ヘテロ接合性 EYS CNV 6:65985254-66085786 ヘテロ接合性	c.2620C>T; p.Gln874Ter バリエーションは、EYS タンパク質の未成熟トランケーションをもたらす未成熟終止コドンを導入します。CNV は、EYS 遺伝子の複数のエクソンを含む約 100 Kb のシーケンスの欠失をもたらします。

まとめ

このユースケースで説明したように、ラボアナリストは TruSight Software Suite の直感的で包括的な機能とツールを使用して、遺伝性疾患に関連するバリエーションを視覚化、トリアージ、解釈できます。

詳細はこちらから

TruSight Software Suite についての詳細は、www.illumina.com/trusight-software-suite.html をご覧ください。

参考文献

- Jaganathan K, Kyriazopoulou Panagiotopoulou S, McRae JF, et al. [Predicting splicing from primary sequence with deep learning](#). *Cell*. 2019;176(3):535–548.
- Sundaram L, Gao H, Padigepati SR, et al. [Predicting the clinical impact of human mutation with deep neural networks](#). *Nat Genet*. 2018;50(8):1161–1170.

TruSight Software Suite について

TruSight Software Suite は、NovaSeq 6000 シーケンスシステムおよび DRAGEN Bio-IT Platform と統合して、NGS の包括的かつ合理化された二次および三次解析ワークフローを可能にする Software as a service (SaaS) 解析ソリューションです。TruSight Software Suite には以下の機能があります。

- **簡単にカスタマイズ可能なケース管理:** シーケンスデータからレポートまでのケースの管理、ユーザーへのケースの割り当て、パイプライン設定の構成、品質管理 (QC) 閾値の設定
- **包括的で超高速なバリエーションコール:** DRAGEN 二次解析を使用して、小さなバリエーション、CNV、構造多型、ミトコンドリアバリエーション、リピート伸長などをコール
- **直感的で強力な解釈:** 遺伝子リスト、遺伝形式、カスタムアノテーション、複雑なロジックを使用したバリエーションのフィルタリングと、重要なバリエーションのフラグ付け、ソート、優先順位付け

現在および将来の機能の一覧については、イルミナの担当営業までお問い合わせください。