

Logiciel Emedgene^{MC}

Optimisez l'interprétation
des variants définie par
l'utilisateur pour les maladies
rares et d'autres applications
de recherche génétique

- Rationalise l'analyse tertiaire pour les applications germinales jusqu'à 75 % à l'aide de l'intelligence artificielle explicable et de l'automatisation personnalisable
- S'intègre aux systèmes LIMS, EMR et à d'autres systèmes informatiques pour simplifier et sécuriser l'ensemble du flux de travail du test
- Permet aux utilisateurs de se développer pour répondre aux demandes en constante évolution de la science, de la technologie et du développement de tests de SNG et de puces à ADN



Introduction

Les applications de séquençage de nouvelle génération (SNG) pour la recherche de maladies rares et d'autres maladies génétiques, la pharmacogénomique, la cytogénétique, d'autres affections germinales ou les études de dépistage des porteurs sont essentielles pour approfondir notre compréhension de la génétique humaine et la convertir en médecine de précision. Le séquençage du génome entier (WGS, Whole-Genome Sequencing), le séquençage de l'exome entier (WES, Whole-Exome Sequencing) et les panels ciblés et virtuels offrent un aperçu authentique et haute résolution de l'ensemble du génome, des régions de codage ou des régions de gènes ciblées, respectivement. Cependant, l'analyse et l'interprétation d'un gros volume de données produites par ces méthodes représentent d'importants goulots d'étranglement pour les laboratoires. C'est le cas plus particulièrement de l'interprétation des variants, un processus manuel chronophage comportant de nombreuses étapes répétitives et prenant jusqu'à sept heures par génome¹. Pour les applications de panel à débit élevé, telles que les projets de dépistage de population saine ou porteuse, les laboratoires cherchent souvent à maximiser l'efficacité et la précision avec des outils d'automatisation, de conservation des variants et de génération efficace de rapports de recherche.

Pour aider les laboratoires à relever ces défis, Illumina propose le logiciel Emedgene, un logiciel de recherche complet conçu par des généticiens pour rationaliser les flux de travail d'interprétation et réduire la charge associée à la conservation des variants. Le logiciel Emedgene est optimisé par l'intelligence artificielle explicable (XAI, Explainable Artificial Intelligence) et l'automatisation pour les flux de travail d'interprétation des variants à débit élevé définie par l'utilisateur, ce qui permet une augmentation de l'efficacité de deux à cinq fois et une réduction de 50 à 75 % de la durée totale du flux de travail par sujet par rapport à l'interprétation manuelle.

En raison de son intégration aux plateformes d'analyse secondaire et de stockage de données d'Illumina, notamment l'analyse secondaire de DRAGEN^{MC}, BaseSpace^{MC} Sequence Hub et Illumina Connected Analytics, le logiciel Emedgene propose des options modulables dans une solution à fournisseur unique, du traitement des échantillons à la génération de rapports de recherche. Le logiciel prend en charge l'entrée de données à partir du WGS, du WES, des panels ciblés et des puces à ADN. Les types de variants complets sont pris en charge, notamment les variants mononucléotidiques (SNV, Single Nucleotide Variant), les insertions/suppressions (indels), les variants du nombre de copies (VNC), les variants de l'ADN mitochondrial (ADNmt), les variants structuraux (SV, Structural Variant) et les répétitions courtes en tandem (STR, Short Tandem Repeats).

Le logiciel Emedgene intègre les capacités complètes d'appel des variants de l'analyse secondaire DRAGEN pour s'assurer que même les gènes et les régions d'intérêt complexes, tels que les variants mosaïques, les régions de duplication segmentaire, etc., sont appelés avec précision comme entrée pour l'interprétation en aval par l'utilisateur final du laboratoire.

Le logiciel Emedgene fournit également l'option de définir des « panels virtuels » par bioinformatique à partir d'une chaîne principale plus complète, telle que le WES ou le WGS, permettant ainsi aux laboratoires de recherche de développer un test en laboratoire de travaux pratiques et de sélectionner le panel de gènes le plus pertinent pour leur application, et rendant les mises à jour du panel de tests plus efficaces à mesure que de nouvelles associations gène-maladie sont découvertes.

Analyse tertiaire complète optimisée par XAI pour tous les types de variants

Les algorithmes d'apprentissage automatique exclusifs fournissent des renseignements appuyés par des preuves et aident les utilisateurs à développer leurs flux de travail d'interprétation. L'IA explicable du logiciel Emedgene met en évidence les variants à examiner en priorité sur la base de preuves conservées associées automatiquement au sein du logiciel. Cela facilite l'élimination des tâches manuelles répétitives qui sont généralement impliquées dans l'association des preuves aux renseignements. L'IA explicable couvre une gamme complète de types de variants, notamment les SNV, les indels, les VNC, les SV, les insertions, les STR et l'ADNmt.

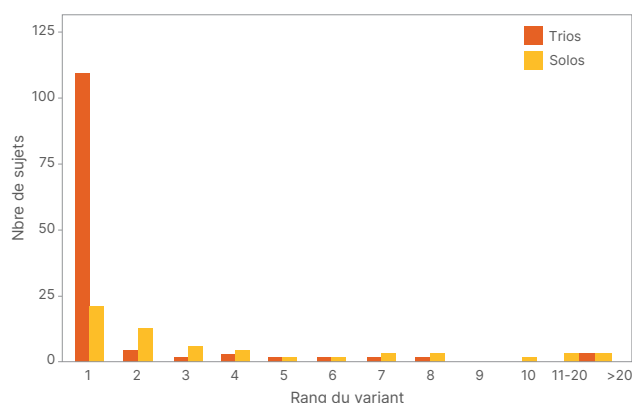


Figure 1 : Classement des variants optimisé par XAI : lors d'une étude de validation menée par des chercheurs de Baylor Genetics, l'interprétation des variants optimisée par XAI a classé correctement les variants causaux dans le top 10 pour 96,6 % des sujets de recherche évalués, notamment des trios (orange) ou des solos (jaune)².

Lors d'une étude de validation, des chercheurs de Baylor Genetics, Houston (Texas) ont évalué l'interprétation des variants optimisée par XAI dans une cohorte de 180 ensembles de données obtenus par WES provenant de sujets de recherche, qui étaient auparavant obtenus par examen manuel. Dans 96,6 % des échantillons évalués, le variant causal a été identifié et classé dans le top 10 pour examen supplémentaire (figure 1)².

Conservation des variants automatisée et explicable

Dans le cadre de son expérience utilisateur rationalisée, le logiciel Emedgene relie les preuves pour chaque variant présélectionné par XAI. Le graphique de preuves reproduit le travail effectué par un scientifique, montrant les associations maladie-gène ainsi que les schémas héréditaires et les phénotypes associés au sujet. Chaque élément de preuve est relié à des sources externes, que ce soit dans la littérature scientifique ou dans une base de données (figure 2). Le graphique de preuves permet un examen et une évaluation rapides des résultats provenant de la plateforme optimisée par XAI.

Les mises à jour mensuelles des sources de données publiques et exclusives garantissent que le logiciel Emedgene dispose des dernières découvertes scientifiques. Le logiciel regroupe et intègre les données provenant de bases de données externes privilégiées, entre autres le catalogue Online Mendelian Inheritance

in Man (OMIM) et ClinVar, dans un graphique de connaissances qui optimise les modèles XAI et inclut la conservation automatique de la littérature scientifique obtenue à l'aide du traitement du langage naturel (TLN).

Classifications automatiques de l'ACMG

Les classifications automatisées de l'American College of Medical Genetics (ACMG) pour les SNV, les indels, les VNC, les SV et les variants d'ADNmt permettent au logiciel Emedgene de rationaliser l'examen des sélections automatisées d'étiquettes de classification effectué par un gestionnaire des variants (figure 3). Par conséquent, la durée d'interprétation des étiquettes quantitatives est réduite et la concordance entre les gestionnaires augmente.

Rationalisation des flux de travail

Le logiciel Emedgene comprend de nombreuses fonctionnalités qui aident à rationaliser et à optimiser les flux de travail d'interprétation définis par l'utilisateur (figure 4). Le logiciel Emedgene est indépendant des tests et compatible avec une large gamme d'entrées de données, notamment les formats de fichiers FASTQ et VCF. Le logiciel s'intègre à la solution de stockage de l'utilisateur pour offrir une expérience sans transfert de données.

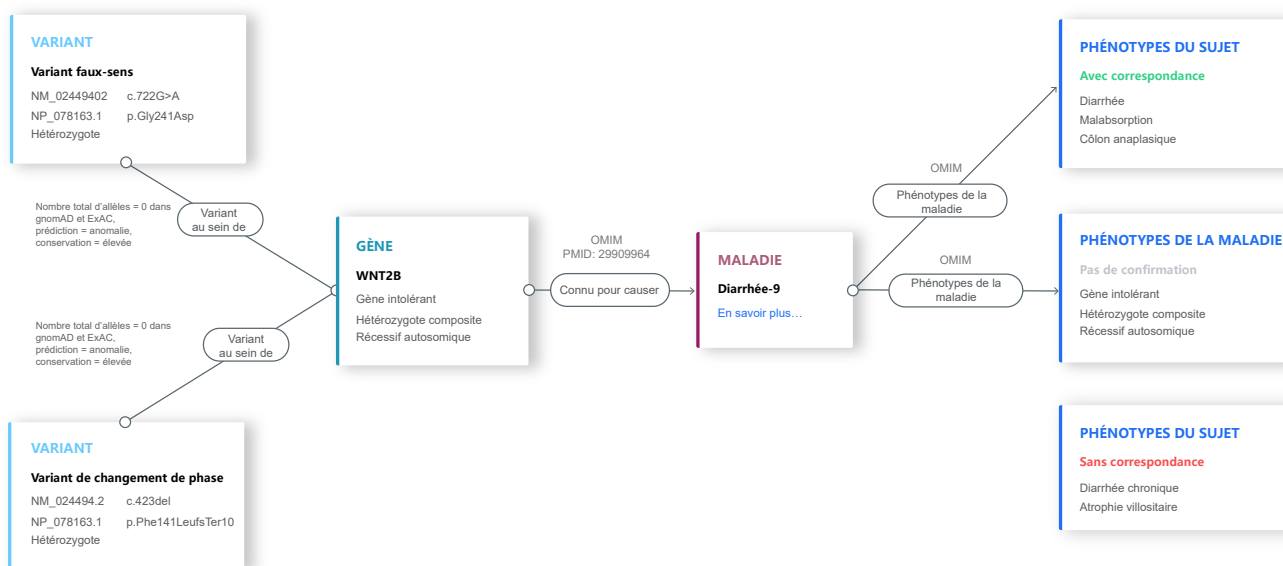


Figure 2 : Graphique de preuves du logiciel Emedgene : les variants bien classés sont indiqués sous la forme d'un graphique de preuves conservées automatiquement compilant les preuves pour un examen rapide et facile par un chercheur en interprétation des variants, y compris des liens vers la littérature et les bases de données pertinentes.

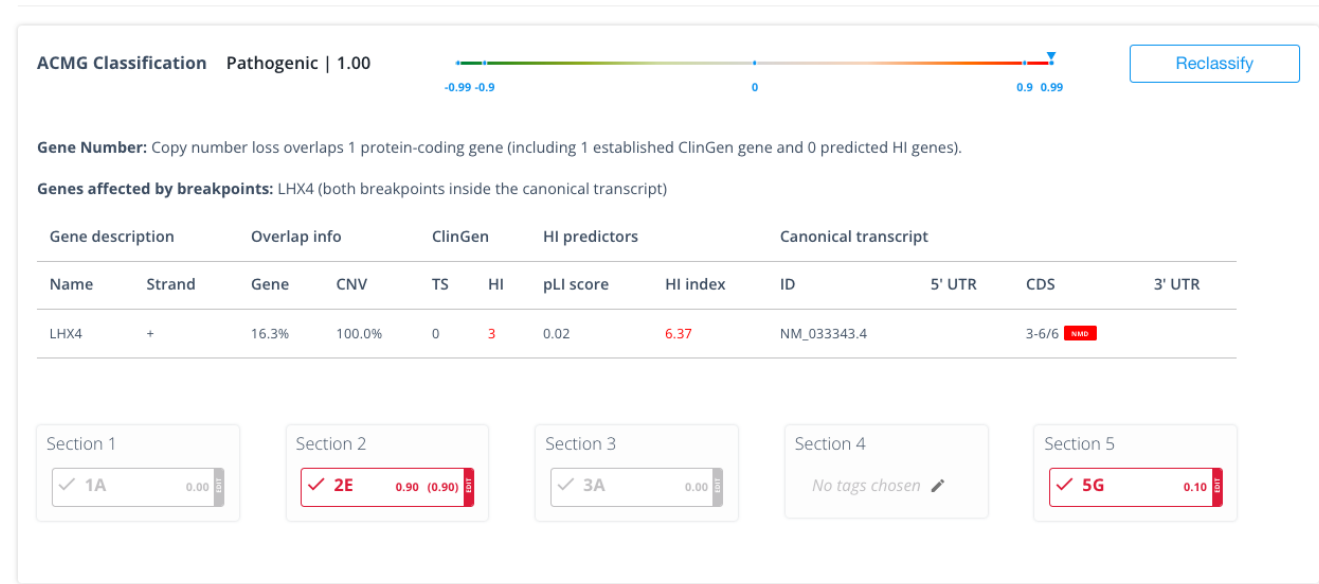


Figure 3 : Classifications automatisées de l'ACMG : le logiciel génère automatiquement les classifications de variants de l'ACMG, notamment les SNV, les indels, les VNC, les SV et les variants d'ADNmt, pour rationaliser l'examen par un chercheur en interprétation des variants.

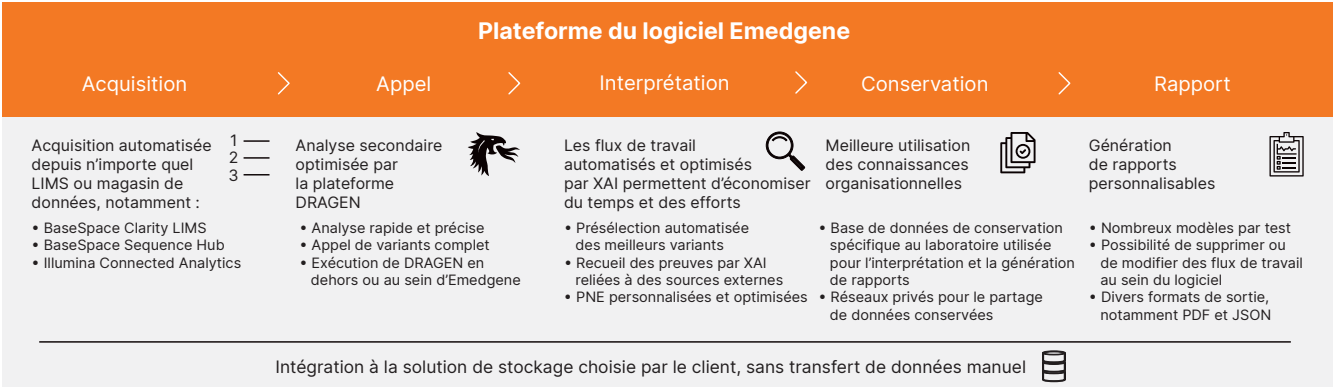


Figure 4 : Le logiciel Emedgene possède de nombreuses fonctionnalités pour rationaliser et optimiser les flux de travail d'interprétation définis par l'utilisateur.

PNE personnalisables

Le logiciel Emedgene offre une plateforme flexible et personnalisable pour permettre aux laboratoires de recherche d'améliorer l'efficacité et de développer leurs flux de travail d'interprétation spécifiques. Des procédures normales d'exploitation (PNE) configurables peuvent être mises en œuvre en créant des ensembles de filtres personnalisés prédéfinis, des flux de travail et des rapports, et par l'inclusion d'environnements de production et de validation.

Conservation spécifique au laboratoire

Le logiciel Emedgene dispose d'un référentiel pour les données conservées d'un laboratoire, qui peut inclure des conservations de variants et de gènes qui sont automatiquement utilisées pour annoter facilement les cas et remplir automatiquement les modèles d'interprétation. La gestion des données conservées et leur constante augmentation se traduisent par des gains d'efficacité importants. La base de données conservée peut faire l'objet d'une recherche par variant et gène, pour un accès facile à tous les cas évalués précédemment.

Partage des données conservées

En tirant parti du pouvoir de la collaboration, les laboratoires peuvent développer leurs connaissances en partageant les données conservées au sein d'un réseau privé de laboratoires. Les partenaires du réseau et les autorisations de partage de données sont détaillés et déterminés par les laboratoires participants. Cette possibilité de partager des données permet aux partenaires de partager des connaissances approfondies avec des partenaires de confiance.

Flux de travail axé sur la visualisation pour l'interprétation des VNC

Des visualisations améliorées à l'échelle du génome et des chromosomes accélèrent l'interprétation des VNC grâce à un accès pratique en un clic aux données au niveau des variants. Les pistes complètes de visualisation de la population et de la base de connaissances sont associées à une visualisation facile des données organisationnelles, historiques et conservées pour fournir une solution rationalisée pour l'interprétation cytogénétique à partir des données de SNG ou de puces à ADN.

Intégration de l'analyse secondaire DRAGEN

Le logiciel d'analyse secondaire DRAGEN est une suite d'applications qui traite les données de SNG et effectue un appel précis, complet et efficace des variants. Dans le cadre du PrecisionFDA Truth Challenge V2, l'analyse secondaire DRAGEN a été nommée comme étant la solution la plus précise (score F1) et la plus

complète parmi les variants évalués dans toutes les régions de comparaison par rapport aux autres outils participants³. La performance éprouvée de la solution permet également de relever les défis courants de l'analyse génomique, tels que le besoin d'une précision optimale de l'appel des variants, la compatibilité avec une large gamme de tests et de types de variants et le traitement efficace de grands ensembles de données génomiques. La référence Multigenome (Graph) capture la diversité génétique, réduit le biais ethnique et améliore la précision des SNV par rapport à d'autres applications. L'analyse secondaire DRAGEN effectue également l'appel complet des variants dans les régions à homologie élevée.

Le logiciel Emedgene fournit l'intégration la plus complète possible avec l'analyse secondaire DRAGEN pour générer une solution d'analyse secondaire à tertiaire optimisée. Les pipelines DRAGEN sont accessibles via une instance locale, BaseSpace Sequence Hub, Illumina Connected Analytics (ICA), ou les pipelines peuvent être configurés et accessibles directement dans le logiciel Emedgene pour une solution prête à l'emploi facilement intégrée. Pour une flexibilité maximale, le logiciel Emedgene est également compatible avec les fichiers de sortie VCF (Variant Call Format) de pratiquement n'importe quelle fonctionnalité d'appel de variant secondaire, ce qui permet aux laboratoires de profiter d'une solution à fournisseur unique pour leurs flux de travail de SNG.

Génération de rapports personnalisables

Les utilisateurs peuvent personnaliser, modifier et remplir automatiquement des rapports en suivant un flux de travail défini par l'utilisateur avec un outil d'édition pratique semblable à Microsoft Word. Grâce à la prise en charge des flux de travail modifiés et révisés, les rapports peuvent être envoyés pour examen supplémentaire au sein du logiciel. Pour davantage de flexibilité et pour faciliter le partage de données, les rapports peuvent être téléchargés en format PDF et JSON.

Intégration avec d'autres plateformes et systèmes

L'intégration des nouveaux instruments ainsi que des nouvelles méthodes et technologies de génomique constitue un défi constant. Le logiciel Emedgene simplifie le processus en s'intégrant aux systèmes NovaSeq[™] X Series, NovaSeq 6000, NextSeq[™] 2000 ou à d'autres systèmes de séquençage d'Illumina pour rationaliser l'analyse du SNG. Le logiciel Emedgene représente également un élément essentiel du flux de travail de recherche sur les maladies génétiques en intégrant une solution « de l'ADN au rapport » aux technologies de préparation de bibliothèques et de séquençage d'Illumina.

Le logiciel Emedgene fournit des interfaces de programmation d'applications (API, Application Programming Interface) qui permettent l'intégration avec les systèmes de gestion des informations de laboratoire (LIMS, Laboratory Information Management Systems) des autres institutions. Des intégrations de stockage sans transfert de données sont disponibles pour BaseSpace Sequence Hub, Illumina Connected Analytics, Amazon S3 pour Amazon Web Services, Azure Blob, et bien plus encore pour une intégration flexible à l'infrastructure informatique existante.

Réseau privé sécurisé

Le logiciel Emedgene respecte les exigences de la loi américaine sur l'assurance maladie (HIPAA, Health Insurance Portability and Accountability Act) et suit les principes du Règlement général sur la protection des données (RGPD) de l'Union européenne. Illumina met en œuvre les principes de protection des données du RGPD, qui exige des organisations qu'elles mettent en œuvre des mesures techniques et organisationnelles pour la sécurité des données et intègrent des mesures de protection aux activités de traitement des données afin de protéger les droits individuels, dès la conception et par défaut. De plus, Illumina offre plusieurs fonctionnalités pour respecter les exigences de la loi HIPAA, notamment des mesures de protection administratives, physiques et techniques.

Le logiciel Emedgene est également conforme aux exigences mondiales et locales en matière de confidentialité des données, notamment la norme ISO 27001, et aux normes SOC 1 et 2 de type II (Service Organization Control). Les options de sécurité avancées personnalisées comprennent l'intégration à la politique d'authentification unique de l'organisation. Une option permettant d'apporter sa propre clé et de doubler le chiffrement des données est également disponible.

Résumé

Le logiciel Emedgene offre une solution basée sur l'automatisation et optimisée par XAI pour l'analyse tertiaire des données de recherche sur les maladies rares et génétiques. Cette solution améliore l'efficacité de l'interprétation des variants et permet aux laboratoires de développer leurs activités. Grâce à son intégration aux systèmes de séquençage, de puce à ADN et logiciels d'Illumina, notamment dans le cadre de l'analyse secondaire DRAGEN pour un appel des variants précis, complet et efficace, le logiciel Emedgene offre aux laboratoires de recherche un outil puissant pour se développer de manière continue.

En savoir plus

[Logiciel Emedgene](#)

Renseignements relatifs à la commande

Produit	N° de référence
Emedgene software – Genome Equivalent Sample (FASTQ)	20073772
Emedgene software – Genome Equivalent Sample (VCF)	20073773
Emedgene software – Annual Support and Maintenance (Basic)	20073776
Emedgene software – Annual Support and Maintenance (Professional)	20073777
Emedgene software – Annual Support and Maintenance (Enterprise)	20073778
Emedgene software – Training at Customer Site	20073774
Custom report (Annual)	20073778
Tous les éléments supplémentaires requis pour le laboratoire seront inclus dans le devis établi.	

Pour les demandes admissibles, Illumina propose une expérience d'évaluation prise en charge, permettant aux clients de travailler avec des exemples d'ensembles de données de recherche disponibles au sein du logiciel Emedgene ou de télécharger et d'évaluer leurs propres ensembles de données au sein du logiciel. Communiquez avec un représentant commercial d'Illumina pour obtenir de plus amples renseignements.

Références

1. Austin-Tse CA, Jobanputra V, Perry DL, et al. [Best practices for the interpretation and reporting of clinical whole genome sequencing](#). *NPJ Genom Med*. 2022;7:27. doi.org/10.1038/s41525-022-00295-z.
2. Meng L, Attali R, Talmy T, et al. [Evaluation of an automated genome interpretation model for rare disease routinely used in a clinical genetic laboratory](#). *Genet Med*. 2023;25(6):100830. doi:10.1016/j.gim.2023.100830
3. Olson ND, Wagner J, McDaniel J, et al. [PrecisionFDA Truth Challenge V2: Calling variants from short and long reads in difficult-to-map regions](#). *Cell Genomics*. 2022;2(5):100129. doi.org/10.1016/j.xgen.2022.100129.



Numéro sans frais aux États-Unis : + (1) 800 809-4566 | Téléphone : + (1) 858 202-4566
techsupport@illumina.com | www.illumina.com

© 2024 Illumina, Inc. Tous droits réservés. Toutes les marques de commerce sont la propriété d'Illumina, Inc. ou de leurs détenteurs respectifs. Pour obtenir des renseignements sur les marques de commerce, consultez la page www.illumina.com/company/legal.html.
M-GL-01057 FRA v2.0