



Notice d'utilisation du logiciel GENEXPATH GenerateReports

Précautions d'utilisation.

RUO

Ce produit est réservé à un usage de recherche dans le domaine de la biologie moléculaire. Ne pas utiliser pour des procédures de diagnostic médical, de prévention ou de traitement.

Il est réservé à un usage professionnel.

Prendre connaissance de l'ensemble des informations portées sur la présente notice avant utilisation.

Contacts :

Fabricant : GENEXPATH

+33 (0)2.78.08.98.69

113 avenue des Martyrs de la Résistance

76100 Rouen - France

contact@genexpath.com

support@genexpath.com



Ce document et son contenu sont la propriété de GENEXPATH. Ils ne sont destinés qu'à l'usage contractuel des clients pour l'utilisation du produit décrit ici à l'exclusion de toute autre utilisation. Ce document et son contenu ne peuvent être utilisés ou distribués dans tout autre but ni communiqués, divulgués ou reproduits de quelque façon que ce soit sans l'accord préalable écrit de GENEXPATH. GENEXPATH ne consent aucune licence à travers ce document. Les instructions données dans ce document doivent être strictement et explicitement suivies par du personnel qualifié et formé afin d'assurer l'utilisation appropriée du produit décrit ici. L'ensemble du contenu de ce document doit être entièrement lu et compris avant l'utilisation du produit décrit.

© 2022 Genexpath. Tous droits réservés.



Table des matières

Préambule – à lire attentivement	6
A propos de ce manuel	7
Description du produit.....	8
Configuration requise	9
Fonctionnalités générales.....	10
Identification et accès à la base de données	10
Fonctionnalités des tableaux	10
Préférences	11
La page d'accueil.....	13
Consultation des données	14
Consulter la liste des rapports de séquençage	14
Consulter les données sur une cohorte d'échantillons	15
Consulter les données sur une série	16
Consulter les données sur une liste créée	17
Consulter les données sur un échantillon.....	18
Consulter les données sur un gène.....	19
Consulter les données sur un variant	19
Données sur les validations externes	19
La page d'accès aux données	21
Identification	21
Données sur les variants	21
Données sur les CNV	23
Ajout de données.....	25
Ajouter un nouveau run PGM.....	25
Ajouter un nouveau run Illumina.....	25
Ajouter des validations externes	26
Prévisualisation des données.....	27



Gestion des Tags de variants	28
La gestion des listes de variants se fait par le menu suivant :	28
Mes tags	28
Mes variants tagués	29
Gestion des listes de variants	30
Visualiser les listes existantes	30
Créer une liste de variants	30
Ajouter des variants à une liste	31
Créer une liste à partir de listes préexistantes	32
La gestion des cohortes	33
Visualisation des cohortes existantes	33
Création d'une nouvelle cohorte	33
Ajouter des échantillons à une cohorte préexistante	34
La gestion des séries	35
Visualisation des séries existantes	35
Création d'une nouvelle série	35
FunEVA	36
Analyses automatiques	37
Aller plus loin avec l'analyse de couverture	37
Comparer la couverture entre les échantillons	38
Aller plus loin avec les séries	39
Ce module bioinformatique permet d'analyser et comparer les résultats obtenus lors des différents suivis pour chaque variation de chaque patient d'une série	39
Aller plus loin avec les listes de variants	40
Log-ratios CNV	42
Ce module permet en renseignant une cellularité estimée d'obtenir un log-ratio théorique en fonction du nombre de copies de segment	42
Les statistiques	44
Statistiques biologiques	44
Les Statistiques du logiciel	44



Distribution des variants.....	44
Les validations externes.....	45



Préambule – à lire attentivement

GENEXPATH GenerateReports est un logiciel bioinformatique. Il doit être utilisé par des professionnels de génétique humaine et avec un esprit critique. Bien que GENEXPATH s'engage à assurer un haut niveau de qualité de ce logiciel, elle ne peut garantir l'exactitude des informations, des annotations publiques et des prédictions qu'il fournit. GENEXPATH se réserve le droit de restreindre les accès au cas de partage des accès non autorisés.



A propos de ce manuel

Ce manuel décrit comment utiliser GENEXPATH GenerateReports version 5.4.2a.



Description du produit

Le logiciel GenerateReports est une plateforme bioinformatique d'interprétation des résultats de séquençage à haut-débit. Celle-ci permet, à partir des données brutes des séquenceurs de nouvelle génération, de créer automatiquement des rapports d'analyse intégrant les données d'identification des échantillons, de qualité ainsi que les listes des variations génétiques ponctuelles et des remaniements du nombre de copies de gène détectés. Chaque instance bénéficie d'une base de données propre permettant une analyse rétrospective des données de séquençage via la constitution de cohortes rétrospectives d'échantillons.



Configuration requise

GenerateReports est une solution bioinformatique disponible à l'adresse [X](#). Son accès est restreint aux comptes utilisateurs Genexpath. Pour vous connecter, vous devez donc posséder un compte et vous authentifier sur l'espace client Genexpath avant de pouvoir consulter les données.

Si vous ne possédez pas de compte, et que vous souhaitez en obtenir un, vous pouvez nous contacter à l'adresse mail suivante : contact@genexpath.com.

La compatibilité du logiciel est assurée sur les navigateurs modernes (Mozilla Firefox, Microsoft Edge, Google Chrome). Il ne nécessite pas d'installation d'outils supplémentaires.

Fonctionnalités générales

Identification et accès à la base de données

Les accès à la base de données sont restreints aux comptes utilisateurs Genexpath.

Si vous ne possédez pas de compte, et que vous souhaitez en obtenir un, vous pouvez nous contacter à l'adresse mail suivante : contact@genexpath.com.

Fonctionnalités des tableaux

Les tableaux générés par la base de données possèdent tous des fonctionnalités communes. Vous pouvez les personnaliser et les exporter.

- [Système de pagination](#)

Un système de pagination est accessible sur tous les tableaux en bas à droite du tableau. Vous pouvez ainsi parcourir les données en utilisant les boutons « previous » et « next ». Il est possible également d'accéder directement à une page en cliquant sur le numéro de page correspondant.

- [Tri des données](#)

Vous pouvez trier par ordre croissant/décroissant ou alphabétique les données numériques/textuelles en cliquant sur les entêtes des tableaux.

Cette fonctionnalité peut être combinée à toutes celles préexistantes. Vous pouvez, par exemple, trier par ordre alphabétique le résultat d'une recherche « Search ». Le tri est aussi compatible avec le système de pagination.

- [Recherche dans les tableaux](#)

Vous pouvez effectuer une recherche textuelle dans tous les tableaux de la base de données en remplissant le formulaire « Search ».

La recherche dans les tableaux peut être combinée avec toutes les fonctionnalités précédemment décrites.

La recherche n'est pas sensible à la casse, c'est-à-dire que le système ne fait pas la différence entre « CHR7 » et « chr7 » par exemple, si vous recherchez toutes les variations sur le chromosome 7.

Attention, le système ne supporte pas la recherche multiple et les mots-clefs de recherche (AND/ET, OR/OU ...).

- [Exportation des tableaux](#)

Chacun des tableaux est exportable en cliquant sur le bouton Excel, CSV, PDF.

Il est également possible d'effectuer une recherche dynamique dans les tableaux par l'intermédiaire du champ de recherche « Search » et de copier uniquement les résultats de la recherche affichés par le bouton « Copy ».

- [L'accessibilité aux données](#)

Il est possible de déterminer la visibilité des informations sur :

- un échantillon en état :
 - visible True : tous les utilisateurs de l'instance ont accès aux informations sur l'échantillon
 - visible False : l'échantillon est anonymisé
- une cohorte en état :
 - Public : tous les utilisateurs de l'instance ont accès aux informations de la cohorte
 - Private : seul l'utilisateur ayant créé la cohorte y a accès

Chaque échantillon doit obligatoirement être associé à au moins une cohorte public.

Préférences

Dans l'onglet « préférences », vous pouvez

- [Configurer les données visualisées](#)

Dans le tableau des variants, vous pouvez configurer les données visualisées en sélectionner différents types de renseignement suivant :

dbSNP	Identifiant de la banque dbSNP (variants) si existant, pointant vers la banque Ensembl
COSMIC_ID	Identifiant de la banque COSMIC (mutations somatiques) si existant, pointant vers la banque COSMIC
COSMIC_tissue	Nombre et tissu dans lequel est retrouvé la mutation, associé au COMIC_ID
ClinVar	Mot clef et identifiant associés à la variation de la banque ClinVar, si existant, pointant vers ClinVar
Strandbias	Pourcentage de biais de brin (favorisation d'un brin par rapport à l'autre)
Homopolymer	Nombre d'homopolymères associé à la variation
n_UMI_wt	Nombre d'UMI sauvages
n_UMI_mt	Nombre d'UMI mutés
n_UMI_discordant	Nombre d'UMI discordants
PopFreqMax	Fréquence maximale des allèles à partir de plusieurs bases de données de fréquence de population
1000G	Fréquence de la variation dans la banque 1000 génomes (ALL : tous, AFR : Afrique, AMR : Amérique, EAS : Asie de l'est, EUR : Europe, SAS : Asie du sud)
ExAC	Fréquence de la variation dans la banque ExAC (ALL : tous, AFR : Afrique, AMR: Amérique, EAS : Asie de l'Est, FIN : Finlande, NFE : Europe hors Finlande, OTH : autres, SAS : Asie du Sud)
ESP6500siv2	Fréquence de la variation dans la banque ESP6500siv2 (ALL : tous, AA : américain d'Afrique , EA : américain d'Europe)
CG46	Fréquence de la variation dans la banque CG46

ADA_score	Score d'AdaBoost de comment le variant peut affecter le splicing
RF_score	Score de Ramdom Forest de comment le variant peut affecter le splicing
Score SIFT	Score de l'algorithmme de prédiction de pathogénicité SIFT (substitution)
Score PolyPhen	Score de l'algorithmme de prédiction de pathogénicité PolyPhen
LRT	Score de l'algorithmme de prédiction LRT (likelihood ratio test) basé sur l'analyse comparative de l'alignement de 32 vertébrés qui permet l'identification de variations impactant des amino-acides conservés au cours de l'évolution
MutationTaster	Score de l'algorithmme de prédiction de pathogénicité Mutationtaster (insertions et délétions)
MutationAssessor	Score de l'algorithmme de prédiction de pathogénicité MutationAssessor (substitutions)
FATHMM	Score de l'algorithmme de prédiction de pathogénicité FATHMM (non synonymous variant and non-coding variants)
PROVEAN	Score de l'algorithmme de prédiction de pathogénicité PROVEAN (amino acid substitution or indel)
VEST3	Score de l'algorithmme de prédiction de pathogénicité VEST3 (missense mutations)
MetaSVM	Score de l'algorithmme de prédiction de pathogénicité MetaSVM
MetaLR	Score de l'algorithmme de prédiction de pathogénicité MEtaLR
MCAP	Score de l'algorithmme de prédiction de pathogénicité M-CAP (rare missense)
CADD	Score de l'algorithmme de prédiction de pathogénicité CADD (insertions et délétions)
fathmm_MKL_coding	Score de l'algorithmme de prédiction de pathogénicité fathmm MLK coding (variant codant et non-codants)
GERP	Score de l'algorithmme de prédiction de pathogénicité GERP
Interpro_domain	Score de l'algorithmme de prédiction des domaines impactés Interpro
Coverage	Couverture de séquençage (Original, +, -, Allele +, Allele -)

- [Configurer la nomenclature des variants par défaut ou selon HGVS.](#)

Cliquer sur « update » pour que les modifications soient prises en compte.

- [Configurer un transcrit favori pour un gène](#)

Vous pouvez choisir un transcrit de gène favori. Les coordonnées seront mises à jour dans les tableaux de variants durant la visualisation des résultats.

Par défaut, tous les transcrits des gènes sont sélectionnés (Full dans le menu déroulant). Pour sélectionner un transcrit, cliquer sur le menu déroulant dans la colonne « choose transcript ».

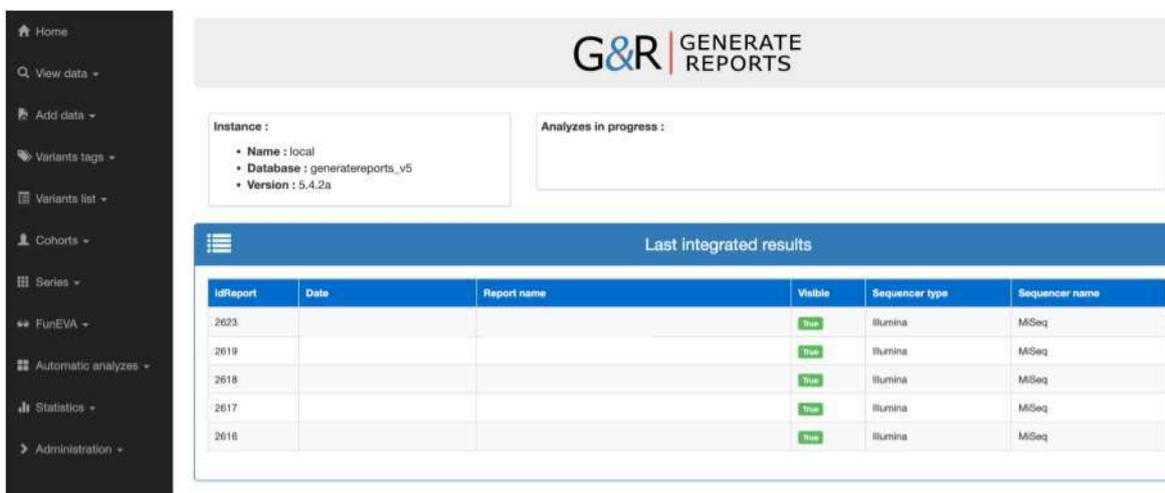
La page d'accueil

La page d'accueil permet d'accéder aux différentes fonctionnalités de l'outil bioinformatique dans le menu de gauche:

- Le retour à la page d'accueil : Home
- La visualisation des données : View data
- L'ajout de données : Add data
- L'ajout de tags sur des variants : Variants tags
- La création et l'utilisation de listes de variants : Variants list
- La création et l'utilisation de cohortes : Cohorts
- La création et l'utilisation de séries : Series
- La recherche d'informations générales sur un gène : FunEVA
- Allez plus loin dans l'analyse des résultats : Automatic analyzes
- Des statistiques : Statistics

En haut à droite, en cliquant sur votre identifiant, vous pouvez accéder :

- Aux préférences : settings
- La déconnexion : log out



The screenshot shows the Genexpath GenerateReports dashboard. On the left is a dark sidebar menu with options: Home, View data, Add data, Variants tags, Variants list, Cohorts, Series, FunEVA, Automatic analyzes, Statistics, and Administration. The main content area has a header with the G&R logo and 'GENERATE REPORTS'. Below the header, there are two boxes: 'Instance' showing details like Name (local), Database (generatereports_v5), and Version (5.4.2a); and 'Analyzes in progress' which is currently empty. The main section is titled 'Last integrated results' and contains a table with the following data:

IdReport	Date	Report name	Visible	Sequencer type	Sequencer name
2623			True	Illumina	MiSeq
2619			True	Illumina	MiSeq
2618			True	Illumina	MiSeq
2617			True	Illumina	MiSeq
2616			True	Illumina	MiSeq

La page d'accueil permet également de visualiser deux types d'informations :

- En haut de l'écran, un rappel sur l'instance auquel vous êtes associés (nom, base de données, version de GenerateReports) et sur les analyses en cours, le cas échéant.
- En bas de l'écran, les derniers résultats intégrés « last integrated results » selon la date, le nom du rapport et le séquenceur utilisé

Consultation des données

La consultation des données est possible en cliquant sur l'onglet « View data ».

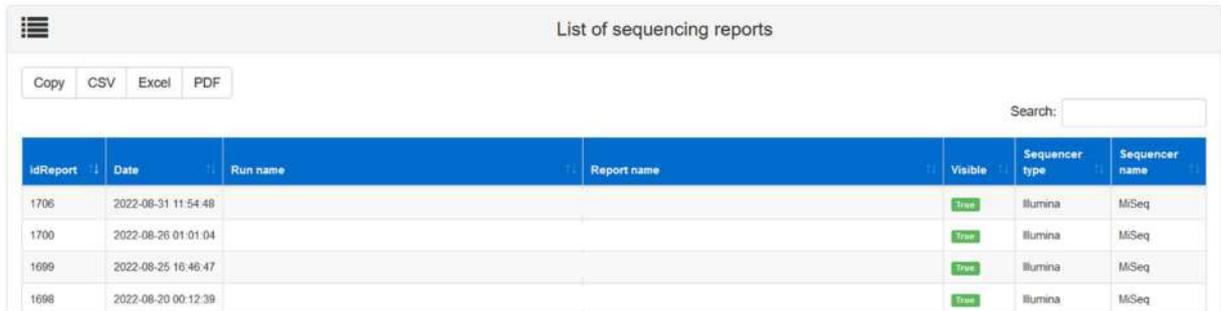


Dans le menu déroulant, différents niveaux de consultation sont possibles en cliquant sur l'onglet correspondant :

- « a report » : la liste des rapports de séquençage
- « a cohort » : la liste des cohortes (ensemble de patients)
- « a series » : la liste des séries (suivi clinique)
- « a list » : la liste des listes de variants
- « a sample » : la liste des échantillons de la base de données
- « a gene » : la liste des gènes séquencés
- « a variant » : la liste des variants séquencés

Consulter la liste des rapports de séquençage

La consultation de la liste des rapports de séquençage permet de récupérer l'ensemble des informations par « run ».



IdReport	Date	Run name	Report name	Visible	Sequencer type	Sequencer name
1706	2022-08-31 11:54:48			True	Illumina	MiSeq
1700	2022-08-26 01:01:04			True	Illumina	MiSeq
1699	2022-08-25 16:46:47			True	Illumina	MiSeq
1698	2022-08-20 00:12:39			True	Illumina	MiSeq

Un tableau affiche la liste des rapports présents dans la base de données de l'instance. Sont également indiqués :

- le numéro d'enregistrement du rapport dans la base de données
- la date de création du rapport,
- le nom du run associé
- le nom du rapport associé
- la visibilité
- le type et le nom du séquenceur utilisé

Ce numéro d'enregistrement permet de distinguer des rapports qui porteraient le même nom dans la base de données.

En cliquant sur le rapport d'intérêt (colonne « report name»), l'ensemble des informations sur ce rapport peuplera ainsi la page de données (voir section *La page d'accès aux données*).

Consulter les données sur une cohorte d'échantillons

La consultation de la liste des rapports de séquençage permet de récupérer l'ensemble des informations sur une cohorte d'échantillons.



idCohort	Cohort name	Samples
88	Public	
86	Private	
85	Private	
84	Public	

Lors de la génération des rapports, les échantillons séquencés sont assignés à une cohorte d'échantillons. Il est possible d'interroger l'ensemble des données associées à une cohorte d'échantillons.

Vous pouvez aussi créer votre propre cohorte d'échantillons biologiques (voir le paragraphe création d'une cohorte).

En cliquant sur le nom de la cohorte d'intérêt, l'ensemble des informations sur le projet sélectionné peuplera ainsi la page de données :



Dans l'encadré de gauche « description », sont renseignés les informations suivantes :

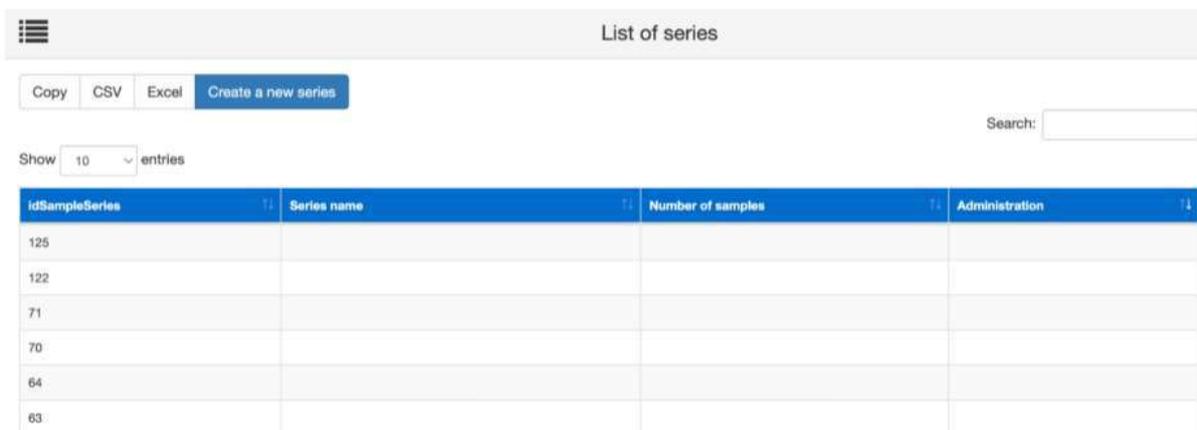
- Le nom de la cohorte
- le nombre d'échantillons, le nombre total de variations ponctuelles (SNV) et le nombre de total de variations de nombre de copies de gène (CNV)
- des informations additionnelles à propos de la cohorte

Dans l'encadré de droite « action », il est possible d'afficher les informations suivantes en cliquant sur le bouton correspondant :

- « cohort settings » permet de modifier le nom de la cohorte et les informations additionnelles
- « list of samples » permet de visualiser l'ensemble des patients de la cohorte et leurs caractéristiques
- « sample statistics » permet d'afficher des informations statistiques sur la cohorte et les patients
- « Show raw variants » permet de visualiser l'ensemble des SNVs et Indels
- « show associated variant list » permet de visualiser, le cas échéant, une liste de variants prédéfinis
- « display all CNV profiles » permet de visualiser, le cas échéant, le profile CNV de chaque patient de la cohorte

Consulter les données sur une série

La consultation des données sur une série permet de récupérer l'ensemble des informations par une série donnée.

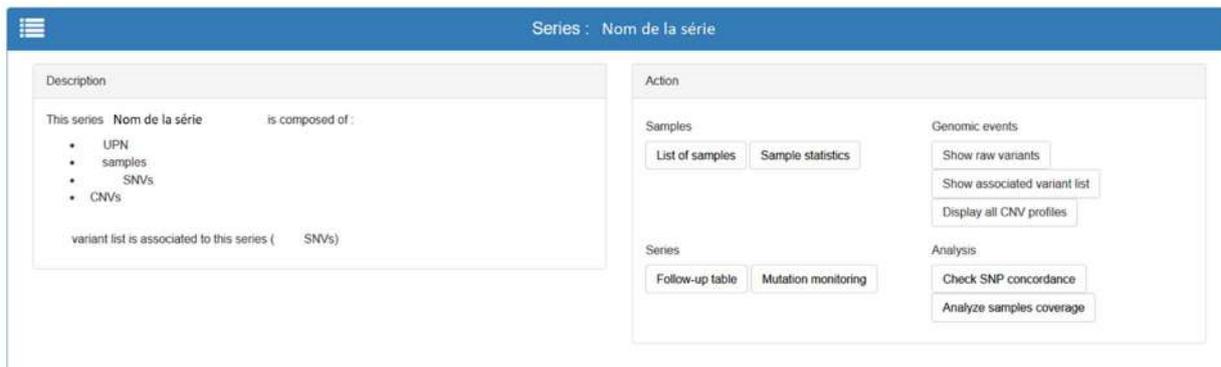


IdSampleSeries	Series name	Number of samples	Administration
125			
122			
71			
70			
64			
63			

La page récapitule l'identifiant de la série, son nom, le nombre d'échantillons de la série et les actions d'administration que l'on peut réaliser sur la série (modifier ou supprimer).

Il est également possible de créer une série à partir de cette page (voir le paragraphe création d'une série).

Pour étudier une série donnée, cliquer sur son nom.



Dans l'encadré de gauche « description », sont renseignés les informations suivantes :

- Le nom de la série
- le nombre d'UPN (unique patient number), d'échantillons, le total de variations ponctuelles (SNV) et le total de variations de nombre de copies de gène (CNV)

Dans l'encadré de droite « action », il est possible d'afficher les informations suivantes en cliquant sur le bouton correspondant :

- « List of samples » : l'ensemble des patients de la cohorte et leurs caractéristiques
- « Sample statistics » : les informations statistiques sur la cohorte et les patients
- « Follow-up table » :
 - o Dans un tableau, l'ensemble des UPN de la série selon l'échantillon correspondant à un évènement clinique unique déclaré pour cette série (Diagnosis, Follow ...). En cliquant sur « compare SNP » dans la colonne « action », la page de comparaison de listes s'ouvre (voir Comparaison de listes de variants)
 - o Dans « series completeness » le suivi de la série selon les évènements cliniques uniques (vert : échantillon présent, rouge : échantillon manquant) ainsi qu'un tableau et un schéma récapitulatifs du nombre d'échantillons selon les évènements
- « Mutation monitoring » :
 - o L'ensemble des mutations associées aux évènements cliniques sont rassemblées dans un tableau. Le nombre de mutations affiché est calculé à partir de la liste des variants associés à cette série.
 - o Une carte des mutations est également disponible
- « Show raw variants » : l'ensemble des SNVs et Indels de la série
- « show associated variant list » : le cas échéant, une liste de variants prédéfinis
- « display all CNV profiles » : le cas échéant, le profil CNV de chaque patient de la cohorte
- « Check SNP concordance » : un tableau récapitulatif des SNP selon les UPN. En cliquant sur « compare SNP » dans la colonne « action », la page de comparaison de listes s'ouvre (voir Comparaison de listes de variants)
- « analyze samples coverage » : les informations sur la profondeur de séquençage de chacune des régions séquencées

Consulter les données sur une liste créée

L'onglet « a list » permet de visualiser la liste des listes créées avec les informations suivantes :

- L'identifiant et le nom de la liste
- Le type de listes

- L'auteur
- Les accès : nom et statuts de l'utilisateur
- Le nombre de SNV, de CNV et d'échantillons uniques de la liste
- Les actions de gestion possibles

Variant list

My variant lists

Show 10 entries

Search:

idVariantList	Name	Type	Date	Author	Accesses	# of SNV	# of CNV	# of unique samples	Action
134		Annotation	2022-09-01 15:06:46		(Editor)	3	0	2	Edit rights Rename Delete
133		Interpretation	2022-09-01 15:00:39			2	0	2	Edit rights Rename Delete

En cliquant sur le nom de la liste d'intérêt, les données de la liste s'affichent en dessous du tableau récapitulatif des listes.

Consulter les données sur un échantillon

La consultation des données sur un échantillon permet de récupérer l'ensemble des informations par « sample name ».

List of samples

Copy CSV Excel

Search:

Sample name	Runs	Visible	Cohorts
		True	

Dans l'onglet « a sample », un tableau de la liste des échantillons présents dans la base de données est affiché ainsi que d'autres informations sur l'échantillon : les runs associés à l'échantillon (liste des séquençages dans lesquels l'échantillon a été séquençé avec son barcode associé), sa visibilité et sa cohorte associée.

Il faut sélectionner l'échantillon d'intérêt dans la colonne « sample name » puis l'ensemble des informations sur l'échantillon sélectionné peuplera ainsi la page de données (voir catégorie page d'accès aux données), quel que soit le nombre de runs associés.

Consulter les données sur un gène

La consultation des données sur un gène permet d'accéder aux données sur tous les runs le concernant.



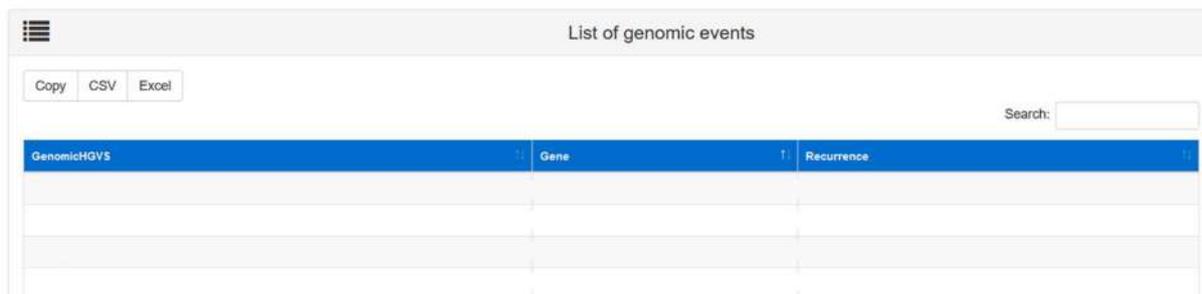
Gene	FunEVA Information	# of variants

Dans l'onglet « a gene », un tableau de la liste des gènes présents dans la base de données est affiché ainsi que d'autres informations sur le gène : la présence d'informations dans FunEVA et le nombre de SNV associés dans toute la base de données.

Il faut sélectionner le gène d'intérêt dans la colonne « gene » puis l'ensemble des informations sur le gène sélectionné peuplera ainsi la page de données (voir catégorie page d'accès aux données), quel que soit le nombre de runs associés.

Consulter les données sur un variant

La consultation des données sur un variant permet d'accéder aux données sur tous les runs le concernant.



GenomicHGVS	Gene	Recurrence

En cliquant sur l'onglet « a variant », un tableau de la liste des variants présents dans la base de données est affiché avec le nom du gène et la récurrence de l'anomalie dans toute la base de données.

En cliquant sur le nom de l'anomalie, l'ensemble des informations sur le variant sélectionné peuplera ainsi la page de données (voir catégorie page d'accès aux données), quel que soit le nombre de runs associé.

Données sur les validations externes

Cette page permet de suivre en temps réel l'ensemble des validations présentes dans la base de données.

☰ List of external validations

Copy CSV Excel PDF

Search:

Genomic HGVS	Sample	Frequency	Positive validations	Negative validations	Result in sample	Matched result	Method	Comment
No data available in table								

Showing 0 to 0 of 0 entries

Previous Next

La consultation des données sur les validations externes permet d'accéder aux données sur la validation par une méthode autre que le NGS d'un variant.

En cliquant sur l'onglet « external validations – Manage validations », un tableau de la liste des variants testés par une autre méthode selon :

- Les coordonnées génomiques HGVS
- Le nom de l'échantillon pour lequel le variant est testé
- La fréquence allélique du variant dans l'échantillon
- Le nombre de validations positives
- Le nombre de validations négatives
- Le résultat rendu pour l'échantillon par la méthode de validation
- Le « matched result » : résultat dans l'échantillon sain du patient
- La méthode utilisée
- Un commentaire

En cliquant sur le numéro d'identification présent dans la première colonne de ce tableau, vous pouvez aussi à tout moment mettre à jour le commentaire concernant la validation. Le commentaire se doit d'être le plus complet possible pour permettre de retrouver plus facilement les raisons d'une saisie particulière de validation.

En cas d'erreur de saisie d'une validation, merci de contacter Genexpath ou l'administrateur de votre instance rapidement pour procéder à la correction.

La page d'accès aux données

La page d'accès aux données permet de consulter et d'importer les données visibles dans les rapports techniques.

On y retrouve 3 sections sous forme de tableau :

- Identification avec des données de qualité du run
- Données sur les variants « SNV & Indels »
- Données sur les CNV « CNV »

Identification

- Dans l'onglet « table », on retrouve un tableau avec une ligne par échantillon. Il permet d'accéder à la liste des échantillons dont les données sont consultées et donne aussi accès :
 - aux rapports techniques au format PDF en cliquant sur le nom de l'échantillon dans la colonne « sample name »
 - aux métriques de qualité par régions ciblées au format CSV en cliquant sur l'item correspondant de la colonne « amplicon depth »
 - à la qualité de l'échantillon, au pourcentage de lectures sur la cible, à l'uniformité et à la profondeur de séquençage.
- Dans l'onglet « quality metrics », des éléments de la qualité de l'échantillon sont disponibles :
 - la qualité de l'échantillon « sample quality », le nombre de lectures séquencées s'alignant sur les régions d'intérêt (« readsOn target ») et l'uniformité de séquençage (« uniformity »). L'uniformité est définie comme étant le pourcentage de bases ciblées ayant au moins $0,2 \times P$ lectures alignées, où P désigne la profondeur moyenne de séquençage.
 - la profondeur de séquençage par région « read depth by region »
- Dans l'onglet « global statistics », les statistiques générales sont disponibles :
 - le nombre de variants par échantillon
 - le nombre de variants par run
 - le nombre d'échantillons par run.

Données sur les variants

- Dans l'onglet « Table », une ligne correspond à un variant et par échantillon avec à minima, par colonne, les informations suivantes:

Action	Possibilité d'ajoutée un variant à une liste de variants avec un commentaire
Lists	Liste(s) de variants contenant le variant
Tags	Tag(s) ajouté(s) au variant

GenerateReports Class	Le type de pathogénicité déterminé par GenerateReports
InterVar	Le type de pathogénicité déterminé par InterVar
Predictive artifact	Cette colonne prédit le caractère artefactuel ou non d'une variation génétique en fonction de la nature des librairies et de l'information portée par les variants. Il indique par exemple, pour les données intégrant des UMI, le ratio UMI concordant/UMI discordant.
Reported in sample	Le variant est présent dans au moins une liste d'interprétation associant celui-ci à l'échantillon (ex : rendu de résultat dans l'échantillon)
Reported	Le variant est présent dans au moins une liste d'interprétation l'associant à un échantillon de la base de données (ex : variant rendu dans un autre échantillon)
Recurrence of validation	Le nombre de fois que le variant a été confirmé ou infirmé par une technique alternative (IGV, Sanger, CGH...).
Recurrence	Le nombre de fois que ce variant a été observé dans toute la base de données de l'instance
Matched validation	Le variant a été confirmé ou invalidé par une technique alternative (IGV, Sanger, CGH...) dans un ADN apparié (ex: constitutionnel)
Sample validation	Le variant a été confirmé ou invalidé par une technique alternative (IGV, Sanger, CGH...) dans l'échantillon visualisé
Gene	Nom du gène/des gènes impacté(s)
Genomic coordinates	Coordonnées génomiques de la variation selon la nomenclature HGVS
Type	Type de variation par rapport à sa conséquence au niveau protéique
Amplicons	Nom de l'amplicon impacté
Quality score	Il s'agit du score de qualité de l'algorithme de Variant Calling utilisé pour générer la liste des variants. Pour les données intégrant des UMI, ce score de confiance varie entre 1000 et 5000. Plus d'informations sur la publication de l'outil UMI-VarCal (https://pubmed.ncbi.nlm.nih.gov/31985795/)
Frequency	Fréquence allélique du variant
UMI Frequency	Fréquence en UMI
Number of mutated reads	Nombre de bases mutées
Number of total reads	Nombre total de bases
Ratio_UMI_disc/mt	Ratio UMI_discordant/UMI_muté
Phased	GenerateReports intègre la notion de phasage de variants. On retrouve dans cette colonne la liste des mutations retrouvées associées à la mutation observée sur un même read.
Phased_Comon_UMI	Nombre d'UMI phasé avec la mutation retrouvées
Phased_UMI_VAF_Ratio	Ratio Phased_UMI/VAF
CLC.singleton_UMI	UMI ayant une seule lecture associée selon la pipeline CLC Genomics
CLC.big_UMI	UMI pour lesquels plusieurs lectures sont associées selon la pipeline CLC Genomics
CLC.prop_singleton	Proportion de singleton selon la pipeline CLC Genomics
CLC.AverageQuality	Qualité moyenne selon la pipeline CLC Genomics

CLC.BaseQRankSum	Comparaison de la qualité de base des données à l'appui de l'allèle de référence avec celles à l'appui de l'allèle alternatif selon la pipeline CLC Genomics
CLC.ReadPositionProb	Probabilité de position du read d'après la pipeline CLC Genomics
CLC.ReadDirectionProb	Probabilité de direction du read selon la pipeline CLC Genomics
Transcripts	Liste des coordonnées cDNA de la variation sur tous les transcrits alternatifs connus du gène, associée aux coordonnées protéiques, selon la nomenclature HGVS
Run	Nom de séquençage associé
report	Nom du rapport
barcode	Nom du barcode associé
Sample	Nom de l'échantillon
Singleton_UMI	Les singletons UMI sont des UMI ayant une seule lecture associée
Big_UMI	Les « big UMI » sont des UMI pour lesquels plusieurs lectures sont associées
Singleton/Big_UMI_ratio	Ratio singleton/big UM
Comments	Commentaires possibles lorsque la variation est associée à une liste de variants

Grâce à l'onglet « préférences », il est possible d'ajouter des colonnes pour obtenir des compléments d'informations (voir sous-partie « préférences »).

En cliquant sur le bouton « column visibility », un menu déroulant apparaît avec les noms de toutes les colonnes du tableau.

- l'onglet « global statistics » affiche des représentations graphiques sur le nombre de variants selon la profondeur de séquençage et le nombre de variant selon la qualité de l'échantillon
- l'onglet « VAF distribution per sample » affiche une présentation graphique de la fréquence de l'allèle muté dans l'échantillon
- l'onglet « Table of occurrence » permet d'obtenir le tableau des occurrences des SNV & Indels

Données sur les CNV

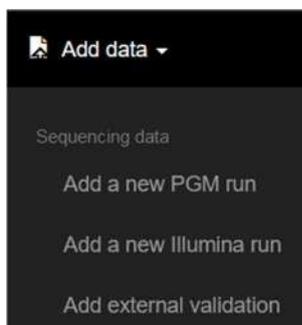
- Dans l'onglet « Table », le tableau correspond à une ligne par CNV et par échantillon, avec :

Lists	Liste(s) de variants contenant le variant
Previously validated ? (somatic)	Précédemment validé
Validation on sample	Nombre de validations déjà répertoriées dans la base de données pour cette variation
Gene	Nom du gène impacté

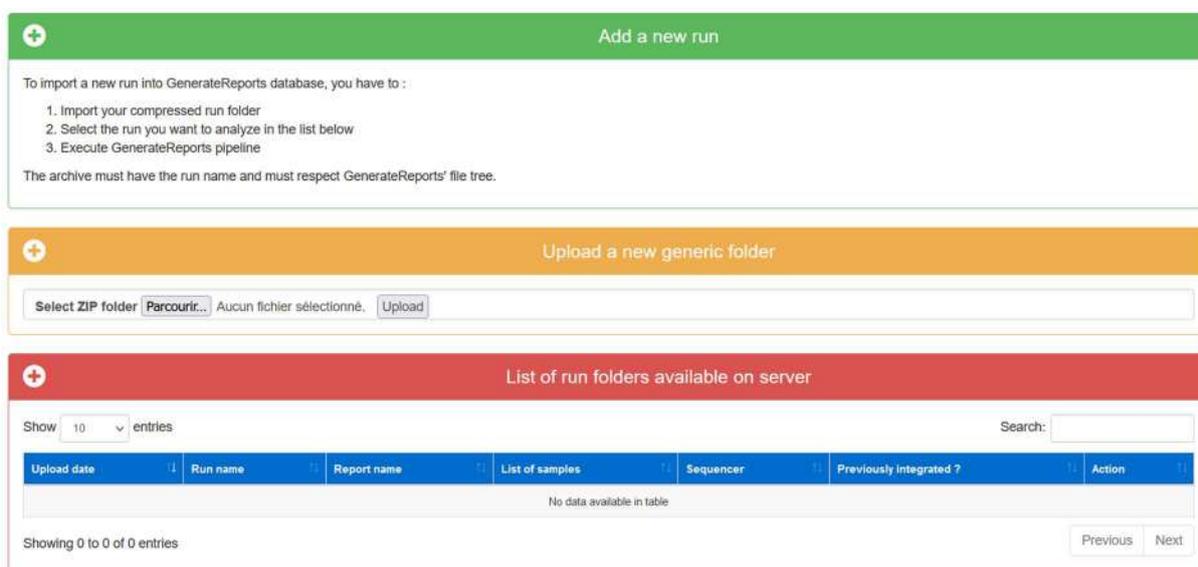
Genomic coordinates	Coordonnées génomiques de la variation telles que prédites par l'algorithme de segmentation DNACopy
Type	Type de CNV (duplication/délétion)
Estimated copy number	Nombre de copies du CNV estimé
P-value	P-value du test comparant les mesures de log-ratio observées dans un segment et la valeur de référence 0.
Q-value	P-value corrigée
Comment	Commentaires éventuels sur l'apparition de points de cassures
Run name	Nom du séquençage
Report name	Nom du rapport
Barcode name	Nom du barcode
Sample name	Nom de l'échantillon

- Dans l'onglet « statistics », la représentation graphique permet de visualiser la répartition des CNV (duplication et délétion) selon les gènes étudiés

Ajout de données



Ajouter un nouveau run PGM



Pour ajouter un nouveau run, il faut :

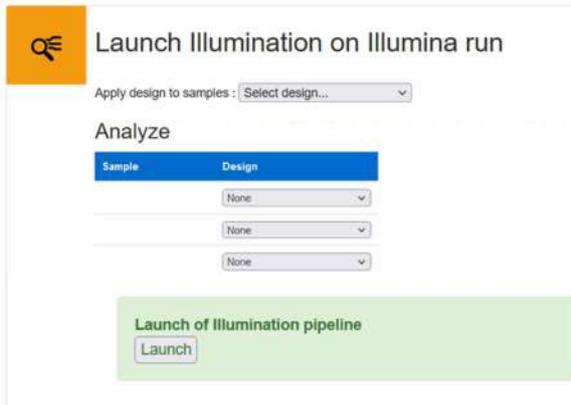
- Importer le fichier du run compressé dans la partie « upload a new generic folder »
- Sélectionner le run à analyser dans la liste « list of run folders available on server »
- Exécuter le pipeline GenerateReport

L'archive doit avoir le nom du run et doit respecter l'arbre de fichiers GenerateReports.

Ajouter un nouveau run Illumina

Pour ajouter un nouveau run, il faut :

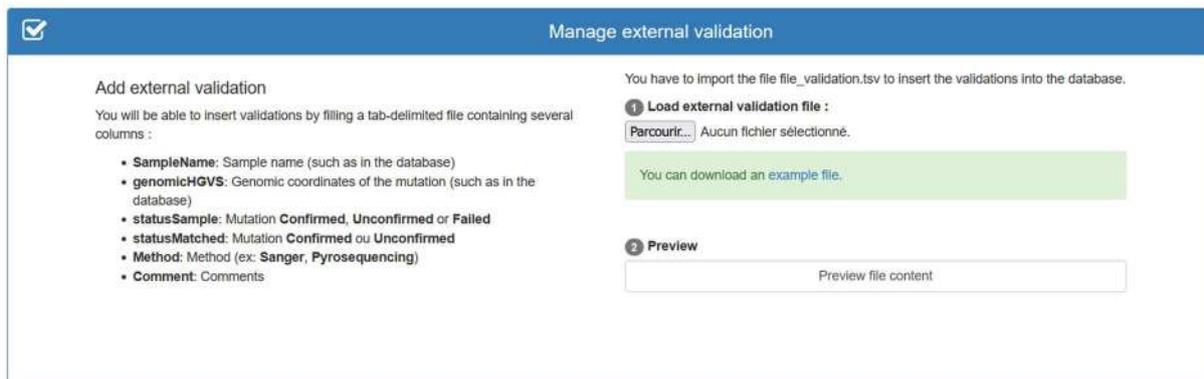
Sélectionner le run d'intérêt en cliquant sur le bouton « Launch »



Une page s'ouvre pour sélectionner les panels de séquençage possibles pour soit tous les échantillons soit pour chaque échantillon.

Appuyez sur le bouton « Launch » pour lancer la (ou les) analyse(s)

Ajouter des validations externes



Pour ajouter des validations, il vous suffit de vous rendre sur cette page et de remplir le fichier d'exemple file_validation.tsv, en respectant un certain nombre de précautions :

- Le nom de l'échantillon doit être exactement le même que dans la base de données
- Les coordonnées de la variation/CNV doivent être exactement les mêmes que dans la base de données
- La colonne statutSample désigne le fait que la variation/CNV ait été validée dans l'échantillon porteur de la variation, ou non. Trois valeurs possibles : Absent, Present, Echec
- La colonne statutMatched désigne le fait que la variation/CNV ait été validée dans un ADN apparié (ex: constitutionnel), ou non. Deux valeurs possibles : Absent ou Présent
- Il est fortement conseillé de mettre un commentaire sur la validation tenant compte des observations. La taille de ce champ est libre.

D'autre part, il n'est pas impossible de rentrer plusieurs validations pour une même variation ou un même CNV si plusieurs méthodes ont été utilisées (Sanger/pyroséquençage...).

Le fichier doit contenir toutes les colonnes du fichier d'exemple. Le séparateur à choisir est une tabulation.



Prévisualisation des données

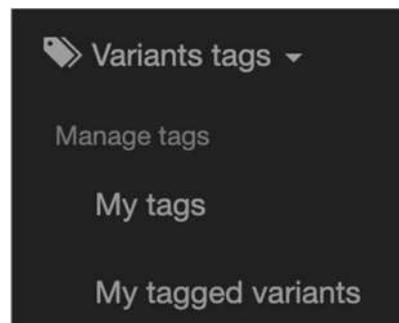
Après l'étape de remplissage du fichier d'exemple, et après l'avoir soumis, une étape de prévisualisation des données à intégrer dans la base de données est réalisée. Cette étape permet de s'assurer du bon déroulement du chargement du fichier, et permet aussi une vérification des données saisies :

- existence de l'échantillon dans la base de données
- existence de l'évènement biologique dans la base de données
- concordance entre l'évènement biologique et l'échantillon dans la base de données
- des duplicatas de validation. Si une validation est déjà présente pour un évènement biologique chez un échantillon donné, faisant intervenir la même méthode d'analyse, alors la ligne du fichier est filtrée. Cela permet d'éviter l'insertion de validations déjà présentes dans la base de données.

Le tableau de prévisualisation tient compte uniquement des données intégrables dans la base de données. Toutes les lignes du fichier comprenant un warning sont filtrés et ne seront pas intégrées dans la base de données.

Gestion des Tags de variants

La gestion des listes de variants se fait par le menu suivant :



Mes tags

Tags

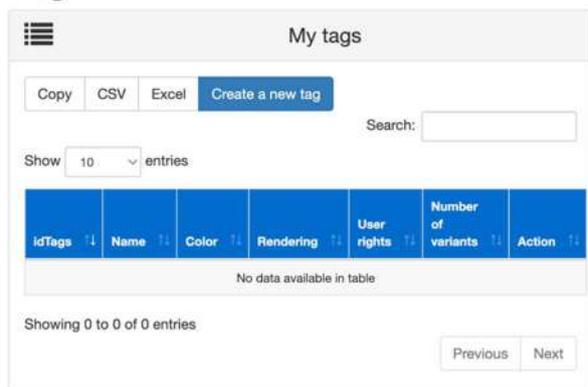


Table interface for 'My tags'. It includes a search bar, a 'Show 10 entries' dropdown, and a table with columns: IdTags, Name, Color, Rendering, User rights, Number of variants, and Action. The table is currently empty, showing 'No data available in table' and 'Showing 0 to 0 of 0 entries'.

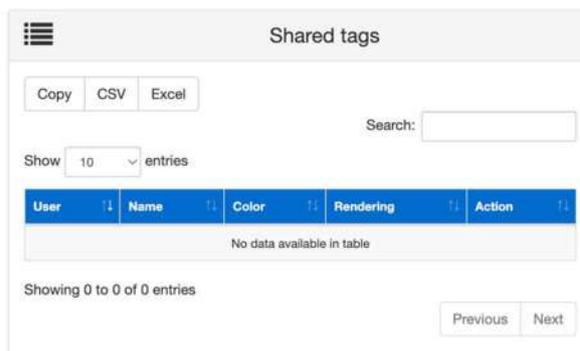


Table interface for 'Shared tags'. It includes a search bar, a 'Show 10 entries' dropdown, and a table with columns: User, Name, Color, Rendering, and Action. The table is currently empty, showing 'No data available in table' and 'Showing 0 to 0 of 0 entries'.

Le tableau de gauche permet de visualiser les tags créés par vous selon :

- L'identifiant du tag
- Son nom
- Sa couleur
- Le rendu de l'affichage du tag
- Les droits d'utilisations du tag
- Le nombre de variant portant ce tag
- Les actions possibles : éditer les droits, renommer le tag et le supprimer

Le tableau de droite permet de visualiser les tags de votre instance selon :

- L'utilisateur qui a créé le tag
- Son nom
- Sa couleur
- Le rendu de l'affichage du tag



- Les actions possibles : éditer les droits, renommer le tag et le supprimer

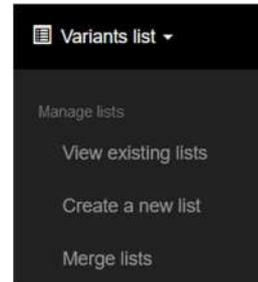
Mes variants tagués

Le tableau permet de visualiser les variants tagués de votre instance selon :

- Identifiant du tag
- Quel tag a été utilisé
- Coordonnées génomiques de la variation selon la nomenclature HGVS
- L'auteur
- Les actions possibles : éditer les droits, renommer le tag et le supprimer

Gestion des listes de variants

La gestion des listes de variants se fait par le menu suivant :



Visualiser les listes existantes

Le tableau permet de visualiser les listes créées par vous ou votre instance selon :

- L'identifiant de la liste
- Son nom
- Le type de listes (annotation ou interpretation)
- La date de création
- L'auteur
- Les droits d'accès
- Le nombre de SNV
- Le nombre de CNV
- Le nombre d'échantillons uniques
- Les actions possibles : éditer les droits, renommer la liste et la supprimer

Créer une liste de variants



Create a list of variants

1 Name :

2 Color :

3 Type : Annotation

You can also create a new list by merging existing list of variants [here](#).

Pour créer une liste, vous devez renseigner dans le formulaire à gauche de la page :

- Son nom
- Lui assigner une couleur : Cette couleur permettra de distinguer

les variants de votre liste d'intérêt dans le tableau des variants en colorant le fond de la ligne avec la couleur choisie si le variant est présent dans la liste que vous venez de définir.

- Son type : annotation pour associer un échantillon et une mutation (colonne Lists du tableau des variants) et interpretation pour associer un échantillon et une mutation et que si la variation génétique ressort à nouveau dans un run, l'information est rapportée (colonne reported dans le tableau des variants)

Puis cliquer sur le bouton « create the list » et sur return une fois la liste créée.

Une liste de variants peut contenir les variants de plusieurs échantillons différents et de plusieurs expérimentations différentes.

Pour créer une liste de variants, vous devez remplir le formulaire.

Une fois la liste créée, vous pouvez lui ajouter des variants.

Ajouter des variants à une liste

Il existe deux façons de procéder pour ajouter un variant à une liste :

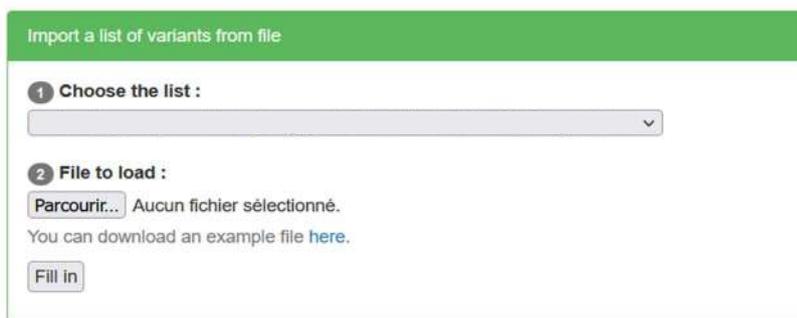
- ajouter le variant directement depuis le tableau des variants

Si votre liste est créée, vous pouvez cliquer sur le bouton de la case List dans le tableau des variants.

Un menu va alors apparaître, vous proposant de sélectionner une liste et éventuellement d'ajouter un commentaire.

Après confirmation de la saisie, le variant se trouve automatiquement ajouté à votre liste.

- en créant un fichier sous excel et en l'ajoutant à la liste



Vous devez absolument partir du fichier d'exemple disponible ici si vous voulez passer par Excel pour ajouter vos variants.

En effet, le fichier d'exemple a été correctement formaté de sorte à pouvoir insérer les

données de manière sécurisée dans la base de données.

Si vous ne passez pas par ce fichier, il est probable qu'un message d'erreur s'affiche lors de la validation du formulaire de saisie.

Le fichier Excel comporte trois colonnes obligatoires :

- Coordonnées génomiques : coordonnées HGVS de la variation, telles que dans la base de données
- Nom de l'échantillon : nom de l'échantillon tel que dans la base de données
- Commentaire : champ libre de saisie

A chaque ligne doit correspondre un variant.

Une fois le fichier créé, il vous suffit de remplir le formulaire de droite.

Vous devez sélectionner la liste dans laquelle les variants seront ajoutés, sélectionner votre fichier, puis valider.

Le logiciel va vérifier l'intégrité des données de votre fichier en plusieurs étapes : le variant est bien présent dans l'échantillon mentionné, et le variant n'est pas déjà présent dans la liste de variants sélectionnés.

Créer une liste à partir de listes préexistantes

Merge lists

You can create a new list using previous lists by merging them.

1 Name of the new list:

2 Select lists to merge :

En remplissant le formulaire de création de listes, vous pouvez créer une nouvelle liste en sélectionnant des listes déjà existantes.

La gestion des cohortes

Visualisation des cohortes existantes

Voir le paragraphe « Consulter les données sur une cohorte d'échantillons »

Création d'une nouvelle cohorte



The screenshot shows a web form titled "Cohort" with a blue header. It contains three main sections:

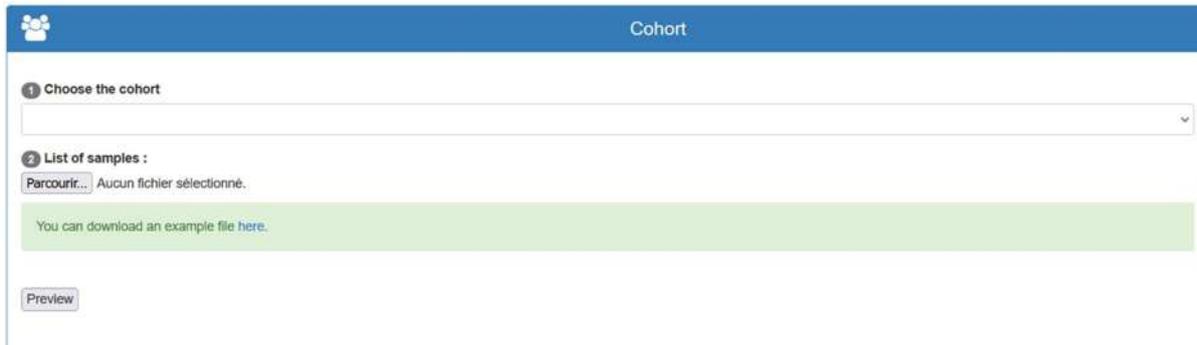
- 1 Cohort name:** A text input field.
- 2 List of samples:** A file selection area with a "Parcourir..." button, the text "Aucun fichier sélectionné.", and a green bar with the text "You can download an example file here."
- 3 Type:** Radio buttons for "Private" (selected) and "Public", and a "Preview" button.

Pour créer votre cohorte, cliquez sur « create a new cohort ». La page « create a new cohort » s'affiche. Il faut:

- Remplir le champs « cohort name » avec le nom de votre cohorte.
- Cliquer sur « parcourir » pour ajouter vos échantillons.
- Sélectionner le type de votre cohorte (private ou public).
- Cliquer sur « Preview » : Une nouvelle page s'ouvre avec un tableau récapitulatif de votre cohorte. Vous pouvez intégrer des données de validation (voir catégorie « insertion des données de validation »).

Si un échantillon a été séquencé plusieurs fois et que vous souhaitez associer ses résultats de séquençage à une cohorte de patients, alors le logiciel vous proposera automatiquement de sélectionner le résultat à retenir pour l'échantillon en question (Assignment).

Ajouter des échantillons à une cohorte préexistante



Grâce à ce formulaire, vous pouvez ajouter des échantillons à une cohorte préexistante en :

- Choissant la cohorte d'intérêt
- Remplissant le fichier d'exemple avec les échantillons souhaités
- Cliquant sur « preview »

La gestion des séries

Visualisation des séries existantes

Voir le paragraphe « Consulter les données sur une série »

Création d'une nouvelle série

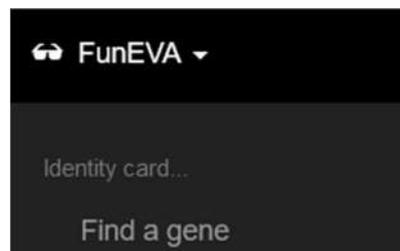
Pour créer votre cohorte, cliquez sur « create a new series ».



La page « create a new cohort » s'affiche. Il faut:

- Remplir le champs « cohort name» avec le nom de votre cohorte.
- Cliquer sur « parcourir » pour ajouter vos échantillons.
- Sélectionner le type de votre cohorte (private ou public).
- Cliquer sur « Preview » : Une nouvelle page s'ouvre avec un tableau récapitulatif de votre cohorte. Vous pouvez intégrer des données de validation (voir catégorie « insertion des données de validation »).

FunEVA



Le plugin FunEVA_gene permet d'obtenir toutes les informations disponibles pour un gène par rapport à la base de données GenerateReports.

Vous obtiendrez des annotations sur les gènes, l'expression, les protéines, les ontologies génétiques, les voies et les interactions.

FunEVA - Gene

Info! FunEVA_gene plugin allows to get all available information for a gene in relation to GenerateReports database. You will obtain annotations on genes, expression, proteins, Gene Ontologies, Pathways and Interactions.

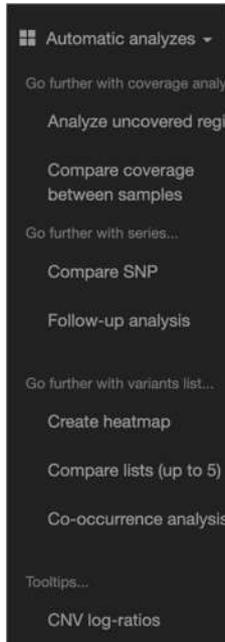
Select a gene

Copy CSV Excel PDF

Search:

Gene name	Old gene names	Aliases	Full name	Ensembl_GeneID	EnTREF_GeneID	CCDS_GeneID	RefSeq_ID	Uniprot_ID
A1BG			alpha-1-B glycoprotein	EH5200000121410	1	CCDS12976	NM_130788	P04217
A1BG-AS1	NCRM00181 A1BGAS A1BG-AS	FLJ23589	A1BG antisense RNA 1	EH5200000026885	50358		NR_015380	
A1CF		ACF ASP ACF4 ACF5 APOBEC1CF	APOBEC1 complementation factor	EH5200000148584	25674	CCDS7241 CCDS7242 CCDS7243 CCDS73133	NM_001198818 NM_014576	Q9NQP4
A2M		FWP907 S8037 CFM025	alpha-2-macroglobulin	EH5200000175899	2	CCDS44927	NM_000214	P01023

Analyses automatiques

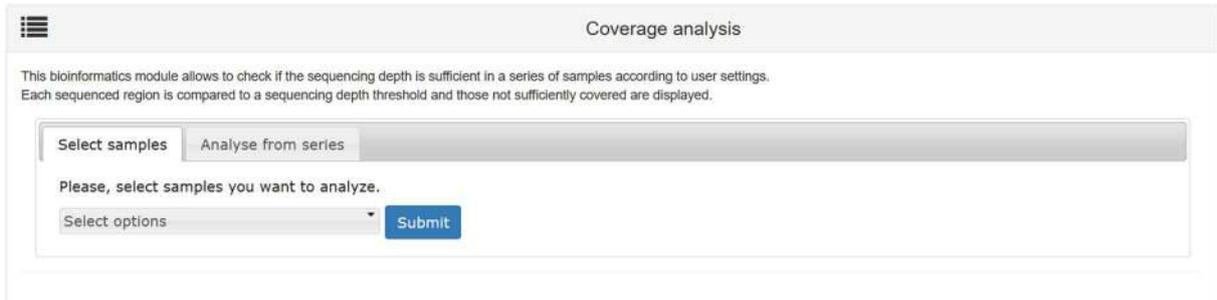


GenerateReports permet également de réaliser des analyses automatiques en fonction :

- De la couverture de séquençage
- Des données de séries
- Des données de listes

Aller plus loin avec l'analyse de couverture

Analyser les régions non couvertes



Ce module bioinformatique permet de vérifier si la profondeur de séquençage est suffisante dans une série d'échantillons en fonction des paramètres de l'utilisateur en les sélectionnant dans une liste d'échantillons (« select samples ») ou à partir d'une série préexistante (« analyse from series »).

Chaque région séquencée est comparée à un seuil de profondeur de séquençage et celles qui ne sont pas suffisamment couvertes sont affichées. Il est possible de modifier les caractéristiques de seuil (en reads ou en UMI) en fonction des contrôles qualité de séquençage du panel utilisé.

Coverage analysis

Settings

Reads		UMI	
Min depth	<input type="text" value="454.53"/>	Min depth	<input type="text" value="261.7"/>
Max depth	<input type="text" value="3029.32"/>	Max depth	<input type="text" value="1234.23"/>

Copy CSV Excel PDF

Search:

idSampleResult	region	chr	start	end	mean read depth	mean umi depth
		chr1	2488029	2488202	2678.21	955.17
		chr1	2489135	2489299	2321.65	886.06

Comparer la couverture entre les échantillons

Analysis settings

Description

This bioinformatics module allows to check if the sequencing depth within a panel is comparable and homogeneous between two or more experimental conditions. The sequencing depth of each region is normalized by sample mean depth to allow the comparison between samples.

Example of use

Let's say 3 samples have been sequenced by two different panel syntheses and we want to compare their coverage. The experimental plan is summarized in the table below :

Group	runName	sampleName
Panel A	231011_NDXS50536_RUO_0256_Example	mySampleNameA
Panel A	231011_NDXS50536_RUO_0256_Example	mySampleNameB
Panel A	231011_NDXS50536_RUO_0256_Example	mySampleNameC
Panel B	231004_NDXS50536_RUO_0255_Example	mySampleNameA
Panel B	231004_NDXS50536_RUO_0255_Example	mySampleNameB
Panel B	231004_NDXS50536_RUO_0255_Example	mySampleNameC

The associated experiment file can be downloaded as an example here.

Launch analysis

Select the file to analyze :

Choisir un fichier

Aucun fichier choisi

Ce module bioinformatique permet de vérifier si la profondeur de séquençage au sein d'un panel est comparable et homogène entre deux ou plusieurs conditions expérimentales.

La profondeur de séquençage de chaque région est normalisée par la profondeur moyenne de l'échantillon pour permettre la comparaison entre deux échantillons.

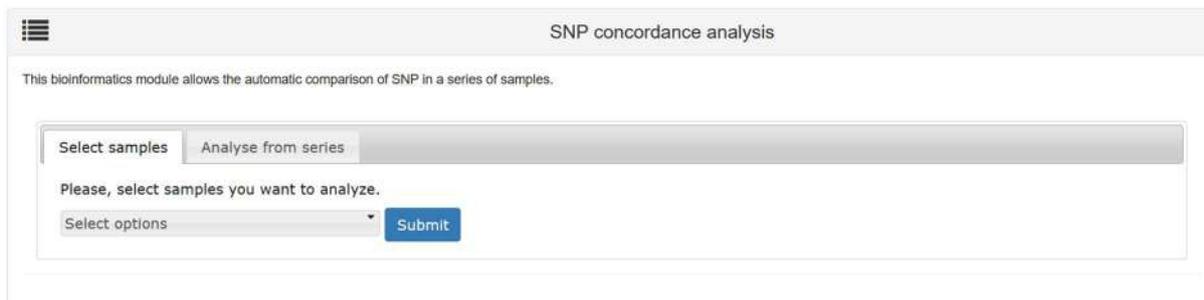
Une analyse peut être lancée en téléchargeant le modèle disponible et en suivant l'exemple d'utilisation :

- On renseigne les groupes d'échantillons dans la colonne « group »

- On renseigne le nom du run où se trouve l'échantillon à analyser dans la colonne « runName »
- On renseigne le nom de l'échantillon à analyser dans la colonne « sampleName »

Aller plus loin avec les séries

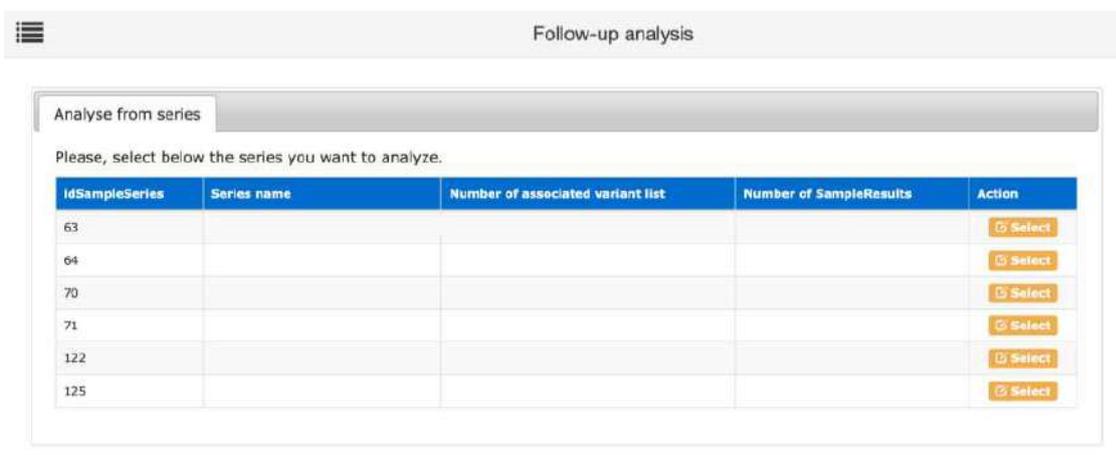
Comparer les SNPs



Ce module bioinformatique permet une comparaison les SNP :

- Entre deux ou plusieurs échantillons par l'onglet « select samples », la page de comparaison de listes s'ouvre (voir le paragraphe Comparaison de listes)
- Entre les échantillons d'une même série par l'onglet « analyse from series », la page d'analyse des concordances de SNP s'ouvre (voir le paragraphe Analyse de concordances de SNP)

Analyse de suivi



IdSampleSeries	Series name	Number of associated variant list	Number of SampleResults	Action
63				Select
64				Select
70				Select
71				Select
122				Select
125				Select

Ce module bioinformatique permet d'analyser et comparer les résultats obtenus lors des différents suivis pour chaque variation de chaque patient d'une série.

L'analyse d'une série peut être lancée en appuyant sur le bouton « Select » dans la colonne « Action ».

Follow-up analysis

Copy Excel

Search:

UPN	Mutation	Gene	Diagnosis	Follow-up
			reads : UMI :	reads : UMI :
			reads : UMI :	reads : UMI :
			reads : UMI :	reads : UMI :

L'analyse donne un tableau renseignant par colonne :

- Le numéro unique du patient
- La mutation
- Le gène où se trouve la mutation
- Les résultats de reads et d'UMI portant la mutation lors du diagnostic
- Les résultats de reads et d'UMI portant la mutation lors du suivi

Aller plus loin avec les listes de variants

- créer une heatmap

Heatmap

Module description

This module allows to aggregate information about CNV and mutations by using an existing variant list.
A heatmap will be created in which genes and samples will be respectively represented by rows and columns.

Data selection

Using an existing variant list...

1 Select the list :

2 Launch

Ce module permet d'agréger des informations à propos des CNV et des mutations en utilisant une liste de variants préexistante.

Pour cela, vous devez :

- sélectionner une liste préexistante
- appuyer sur Launch



La heatmap est générée avec :

- en abscisse les échantillons
- en ordonnée les variants étudiées (par symboles, le numéro de l'altération génétique ou la moyenne de la fréquence mutationnelle)
- les cellules peuvent représenter l'altération génomique, la mutation ou le statut du CNV

Il est possible de télécharger l'image en cliquant sur le bouton correspondant.

- Comparer des listes

☰ Variant list comparison

Compare variant lists

This module allows to create a Venn diagram from sequencing data from several variant lists.

The Venn diagram will allow you to see common or exclusive variants between lists.

The variant lists should be previously created using [this form](#).

Start the comparison

Using variant lists...

1 Select lists :

2 Realize the comparison

Using an external file...

You can download an example file [here](#).

1 Select the file to load :

Aucun fichier sélectionné.

2 Realize the comparison

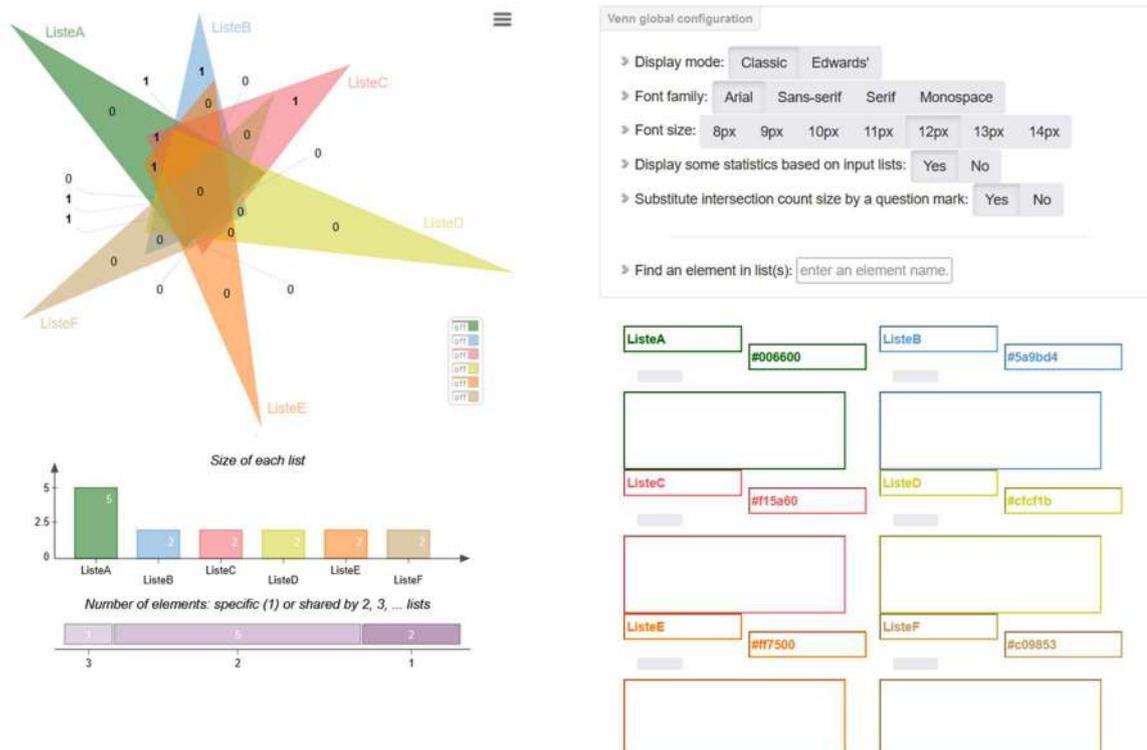
Ce module permet de créer à un diagramme de Venn à partir des données de séquençage de plusieurs listes de variants (jusqu'à 5 listes en simultanée).

Ce type de diagramme permet d'observer les variants communs ou exclusifs entre les listes.

Cette comparaison peut être réalisée :

- En sélectionnant des listes de variants préexistantes et en appuyant sur « Launch »
- En chargeant un fichier modèle avec les listes d'intérêt et en appuyant sur « Launch »

☰ Variant list comparison



En cliquant sur l'intersection de 2 ou plusieurs listes, un tableau apparaît avec différentes informations sur les variants d'intérêt (voir le paragraphe La page d'accès aux données).

Log-ratios CNV

Ce module permet en renseignant une cellularité estimée d'obtenir un log-ratio théorique en fonction du nombre de copies de segment.

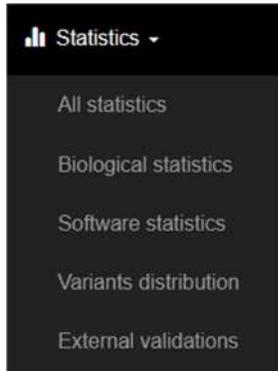


Compute log-ratios from celularity

To compute log-ratios, please fill the measured estimated cellularity :

Segment copies	Theorial log-ratios
5	<input type="text" value="0"/>
4	<input type="text" value="0"/>
3	<input type="text" value="0"/>
2	<input type="text" value="0"/>
1	<input type="text" value="0"/>
0	<input type="text" value="0"/>

Les statistiques



Ce module bioinformatique permet d'étudier les différentes statistiques de la base de données de l'instance :

- Toutes les statistiques
- Les statistiques biologiques
- Les statistiques du logiciel
- Les statistiques sur la distribution des variants
- Les statistiques sur les validations externes

Statistiques biologiques

Cet onglet permet d'observer la répartition de différents paramètres via une représentation graphique :

- Le nombre de variants par échantillon « variants number per sample »
- Le nombre de variants par run « variants number per run »
- Le nombre d'échantillon par run « number of sample per run »
- Le nombre de variants selon la profondeur « variants number = f(depth)
- Le nombre de variants en fonction de la qualité de l'échantillon « variants number = f(sampleQuality) »

Les Statistiques du logiciel

Cet onglet permet d'observer les informations sur :

- La répartition des XXX de la base de données
- La volumétrie des données en fonction des informations listées en légende
- Le monitoring en fonction du temps et de la taille des données

Distribution des variants

Le graphique indique la proportion de chaque type de variants présents dans la base de données GenerateReports.

Le tableau permet d'observer selon le gène d'intérêt le nombre de

- Splicing exonique
- Une délétion induisant un décalage du cadre de lecture
- Une insertion induisant un décalage du cadre de lecture
- Une substitution induisant un décalage du cadre de lecture

- Une mutation intergénique (mutations dans des gènes différents peuvent conduire à la même maladie) ?
- Une mutation intronique
- Non codant RNA exonique
- Non codant RNA intronique
- Une délétion n'induisant pas un décalage du cadre de lecture
- Une insertion n'induisant pas un décalage du cadre de lecture
- Une substitution n'induisant pas un décalage du cadre de lecture
- Un SNV non synonyme
- Un splicing
- Un gain d'un codon stop
- Une perte d'un codon stop
- Un SNV synonyme
- Un upstream
- Une mutation dans la région 3' UTR
- Une mutation dans la région 5' UTR

Les validations externes

Cet onglet permet d'observer :

- les statistiques des validations externes selon le type de technique en fonction du nombre de validations et selon leur état (erreur, non confirmé ou confirmé)
- le score variantCaller en fonction des résultats Sanger.