

Introduction à la bioinformatique (UE SSV3U15)

TP6. Variations génétiques

Diaporama d'accompagnement du TP

Conception : Aitor Gonzalez

Révision : Emese Meglec et [Jacques van Helden](#)

Objectifs

- Apprendre à utiliser des ressources bioinformatique en ligne pour analyser le lien entre variations génétiques et maladies humaines
- Faire le lien entre la diversité inter-populationnelle des variations et la répartition géographique de certaines maladies génétiques

Exemples traités durant le TP

- Au cours de ce TP, nous étudierons une maladie humaine appelée drépanocytose (synonyme : anémie falciforme). Nous explorerons la base génétique de cette maladie et nous nous baserons sur cette analyse génétique pour comprendre les différences de sa prévalence dans différentes populations.

Notions mises en pratique

- Variation génomique, variant
- Relations séquence - structure - fonction d'une protéine
- Bases de données biologiques : maladies, gènes, protéines et variants génétiques

N'oubliez pas que vous pouvez à tout moment consulter le [glossaire du cours](#) pour obtenir une définition sommaire des principaux termes utilisés.

Compétences

A l'issue de ce TP, vous devriez avoir acquis les compétences suivantes.

- Rechercher des informations sur une maladie
- Identifier les populations le plus touchés sur une maladie
- Identifier les gènes associés à maladie génétique
- Rechercher les variations génétiques connues responsables de la drépanocytose
- Analyser la fréquence des variations génétiques dans des populations
- Analyser l'impact des variations génétiques sur la santé
- Faire le lien entre modifications de la séquence/structure/fonction d'une protéine et pathologie associée

Etapes

1. **Medline Plus**, base de données grand public
2. **OMIM**, catalogue professionnel des gènes et maladies génétiques associées
3. **UCSC genome browser** : analyse génomique du gène de l'hémoglobine B et du variant rs334
4. **dbSNP** : fréquence populationnelle des variants
5. **Pubmed** : impact du variant rs334 dans les populations africaines

Complétion

- Tous les exercices doivent être réalisés par chaque étudiant.
- En principe, l'ensemble des exercices devraient être complétés en séance (avec explications par les enseignants).
- Si nécessaire, ils pourront être terminés ultérieurement.

Définition des concepts

Variation génétique	Différence entre les séquences d'ADN des individus d'une population. Note : les variations peuvent résulter de substitutions, de délétions, d'insertions, de duplications ou de translocations. Dans ce TP nous nous intéresserons à des substitutions limitées à un seul nucléotide (SNP).
Variant	Région génomique qui présente des séquences (allèles) différentes entre individus d'une population. Exemple : T ou C à la position 293201 du chromosome 5. Un variant peut avoir deux ou plusieurs allèles (par exemple T ou C ou G).
SNP	Single Nucleotide Polymorphism : variant qui résulte d'une substitution, limitée à seul nucléotide
Allele	L'une des séquences d'une variation. Exemple : C à la position 293201 du chromosome 5.
Allèle majeur	Allèle dont la fréquence est la plus élevée dans une population donnée.
Allèle(s) mineur(s)	Allèle(s) dont la fréquence n'est pas la plus élevée dans une population donnée.
Allèle de référence	Séquence du génome de référence pour un variant donné. Le génome de référence a été séquencé en combinant des fragments de chromosomes issus de différentes personnes (par souci de protection des données personnelles). Les allèles de référence peuvent donc correspondre à l'allèle majeur ou à un allèle mineur, selon les particularités des individus ayant servi pour le séquençage du projet génome humain initial.
Allèle alternatif	Séquence différente de l'allèle de référence pour un variant donné
Mutation	Evènement lié à la réplication de l'ADN, qui suscite l'apparition d'une nouvelle forme de la séquence, transmissible aux générations suivantes. Les variations qu'on observe entre individus d'une population résultent pour la plupart de mutation qui se sont produites chez un individu ancestral.

Ressources bioinformatiques utilisées

Nom	URL	Description
MedlinePlus	https://medlineplus.gov	A trusted health information website that provides easy-to-understand medical information and resources for patients and healthcare professionals.
Centers for Disease Control and Prevention (CDC)	https://www.cdc.gov	The national public health agency in the U.S. responsible for disease prevention, control, and public health data dissemination.
UniProt	https://www.uniprot.org	A comprehensive, freely accessible database of protein sequence and functional information used for biological research.
UCSC Genome Browser	https://genome.ucsc.edu	A web-based tool for exploring the reference genome sequences and their annotations across various species, enabling researchers to visualize genomic data.
NCBI dbSNP	https://www.ncbi.nlm.nih.gov/snp	A database of single nucleotide polymorphisms (SNPs) and other small genetic variations, hosted by the National Center for Biotechnology Information (NCBI).
International Genome Sample Resource (IGSR)	https://www.internationalgenome.org	A repository and resource for genomic data from global populations, building on the 1000 Genomes Project to support genetic diversity research.
PubMed	https://pubmed.ncbi.nlm.nih.gov	A free search engine primarily accessing the MEDLINE database of references and abstracts on life sciences and biomedical topics.

1. Recherche Medline plus : qu'est-ce que la drépanocytose ?

Dans un premier temps nous allons nous renseigner sur la maladie en consultant un site web destiné au grand public, qui est géré par un organisme scientifique public : **Medline Plus**.

MedlinePlus est un site de référence en matière d'informations médicales, offrant des ressources fiables et faciles à comprendre sur de nombreux sujets de santé. Il est géré par la Bibliothèque Nationale de Médecine des États-Unis (NLM), une institution publique de recherche biomédicale.

- Ouvrez une connexion à **Medline plus** (medlineplus.gov)
- Effectuez une recherche de l'expression "**sickle cell disease**" (nom anglais de l'anémie falciforme)
- Lisez le résumé "What- is sickle cell disease (SCD) ?" en cliquant "**Read more**" pour accéder à la fiche complète.

Sur Ametice, répondez au Questions 1 à 5 du TP6.

1. q.01mlp.1: Quelle est la cause principale de la drépanocytose?
2. q.01mlp.2: Parmi les symptômes suivants, cochez ceux qui sont communs dans la drépanocytose.
3. q.01mlp.3: Quelle est la différence entre "sickle cell disease" et "sickle cell trait" ?
4. q.01mlp.4: Comment la drépanocytose est-elle héritée ? Note: pour trouver la réponse, vous devez accéder à une ressource externe à Medline, en suivant le lien "Genetics".
5. q.01mlp.5: Quelles sont les populations les plus touchées par cette maladie aux Etats-Unis ?

The screenshot shows the MedlinePlus website interface. At the top, there is a search bar with the text "sickle cell disease" and a "GO" button. Below the search bar, there are navigation links for "Health Topics", "Drugs & Supplements", "Genetics", "Medical Tests", "Medical Encyclopedia", and "Español". The main content area displays search results for "Sickle Cell Disease". On the left, there are two "Refine by" sections: "Refine by Type" and "Refine by Format". The "Refine by Type" section lists categories such as Health Topics (9), Medical Encyclopedia (53), External Health Links (176), Medical Tests (8), Genetics (13), Drugs and Supplements (25), National Institutes of Health (43), MedlinePlus Magazine (108), Healthy Recipes (0), and Multiple Languages (0). The "Refine by Format" section lists PDF (6), Images (3), and Videos (11). The main search results area shows a snippet for "What is sickle cell disease (SCD)?" with a small image of red blood cells. A red circle highlights the "(Read more)" link. Below the snippet, there is a list of results, with the first result being "1. Sickle Cell Disease (National Library of Medicine)" and the second being "2. Sickle cell disease".

2. Consultation d'OMIM, le catalogue des maladies génétiques humaines

Pour rappel, nous avons vu au cours que l'hémoglobine est un hétérotétramère composé de deux paires de chaînes polypeptidiques encodées respectivement par les gènes HBA (hémoglobine alpha) et HBB (hémoglobine beta).

Nous allons maintenant consulter la base de connaissance OMIM pour identifier le(s) gène(s) dont les variations génétiques sont associées à la drépanocytose. OMIM est un catalogue complet et d'accès public, qui répertorie les gènes humains et les maladies génétiques associées. Il est conçu pour aider les professionnels de la santé, les chercheurs et les étudiants dans les domaines de la génétique et de la médecine. Il nous informe sur les maladies, les gènes et les variants.

- Ouvrez une connexion au site Web OMIM (www.omim.org)
- Effectuez une recherche avec la requête “**sickle cell disease**”.
- Parcourez brièvement la description de la pathologie, et notez les différences avec Medline Plus (niveau de détail, vocabulaire utilisé, références bibliographiques)

Sur Ametice, répondez aux questions suivantes

6. q.02omm.1: Quel est le symbole du gène qui cause la drépanocytose?
7. q.02omm.2: Dans quelle bande cytogénétique trouve-t-on ce gène?

#603903 ICD+
Table of Contents

Title
Phenotype-Gene Relationships
Clinical Synopsis
Text
Description
Clinical Features
Diagnosis
Clinical Management
Pathogenesis
Molecular Genetics
Population Genetics
Animal Model
History
References
Contributors
Creation Date
Edit History

603903

SICKLE CELL DISEASE

Alternative titles; symbols
SICKLE CELL ANEMIA

Phenotype-Gene Relationships

Location	Phenotype	Phenotype MIM number	Inheritance	Phenotype mapping key	Gene/Locus	Gene/Locus MIM number
11p15.4	Sickle cell disease	603903	AR	3	HBB	141900

Clinical Synopsis - PheneGene Graphics +

▼ **TEXT**

A number sign (#) is used with this entry because **sickle cell disease** is the result of mutant beta globin (HBB; 141900) in which the mutation causes sickling of hemoglobin.

▼ **Description**

Sickle cell disease is a multisystem disease associated with episodes of acute illness and progressive organ damage. Hemoglobin polymerization, leading to erythrocyte rigidity and vasoocclusion, is central to the pathophysiology of the disease, but the importance of chronic anemia, hemolysis, and vasculopathy has been established. The most common cause of sickle cell anemia is the HbS variant (141900.0243), with hemoglobin SS disease being most prevalent in Africans (review by Rees et al., 2010). +

See review of infection in sickle cell disease by Booth et al. (2010). +

Piel et al. (2017) reviewed the genetic and nongenetic modifiers of the severity of **sickle cell disease**. +

▼ **Clinical Features**

Scriver and Waugh (1930) reported detailed studies of a 7-year-old child with sickle cell anemia. Her main complaints were cough, night sweats, vague pains in the legs and joints, occasional abdominal pain, poor appetite, and increasing fatigue. In a series of clever experiments that involved taking

2. Consultation d'OMIM, le catalogue des maladies génétiques humaines

Après avoir identifié le gène associé à la drépanocytose, nous voulons rechercher des variants-clés de cette pathologie. Les informations sur les variants se trouvent dans la section “**molecular genetics**” d'OMIM.

Le site OMIM utilise un système interne pour identifier les gènes et les maladies. Par exemple, l'identifiant du gène BRCA1 est 113705. Ensuite, les variants sont référencés par un identifiant qui commence par l'identifiant du gène, suivi d'un point et de quatre chiffres. Par exemple, les variants de BRCA1 sont référencés par 113705.0001, 113705.0002, ...

Sur Ametice, répondez à la question suivante

8. q.02omm.3: Selon la base de données OMIM, quel est le variant génétique causal le plus fréquent pour la drépanocytose ?

#603903

Table of Contents

Title

Phenotype–Gene Relationships

Clinical Synopsis

Text

Description

Clinical Features

Diagnosis

Clinical Management

Pathogenesis

Molecular Genetics

Population Genetics

Animal Model

History

References

Contributors

Creation Date

Edit History

▼ Molecular Genetics

The most common cause of sickle cell anemia is HbS (141900.0243), with hemoglobin SS disease being most prevalent in Africans. Rees et al. (2010) listed genotypes that had been reported to cause sickle cell disease. +

Modifier Genes

Priapism, a vasoocclusive manifestation of sickle cell disease, affects more than 30% of males with the disorder. In sickle cell anemia patients, 148 with priapism and 529 without, Nolan et al. (2005) searched SNPs from 44 genes of different functional classes for an association with priapism. By genotypic and haplotype analysis, they found an association between SNPs in the KLOTHO gene (604824) and priapism (rs2249358 and rs211239; adjusted odds ratio of 2.6 and 1.7, respectively). Nolan et al. (2005) noted that the finding may have broader implications in sickle cell disease, as the KL protein regulates vascular functions, including the expression of VEGF (192240) and release of endothelial nitric oxide (see 163729). +

Sickle cell anemia is phenotypically complex, with different clinical courses ranging from early childhood mortality to a virtually unrecognized condition. Overt stroke is a severe complication affecting 6 to 8% of individuals with sickle cell anemia. Modifier genes might interact to determine the susceptibility to stroke. Using Bayesian networks, Sebastiani et al. (2005) analyzed 108 SNPs in 39 candidate genes in 1,398 individuals with sickle cell anemia. They found that 31 SNPs in 12 genes interacted with fetal hemoglobin to modulate the risk of stroke. This network of interactions included 3 genes in the TGF-beta pathway (see 190180) and SELP (173610). Sebastiani et al. (2005) validated their model in a different population by predicting the occurrence of stroke in 114 individuals with 98.2% accuracy. +

Uda et al. (2008) found that the C allele of rs11886868 in the BCL11A gene (606557.0002) was associated with an ameliorated phenotype in patients with sickle cell anemia, due to increased production of fetal hemoglobin. +

2. Consultation d'OMIM, le catalogue des maladies génétiques humaines

La base de données la plus complète de variations génétiques humaines est **dbSNP du NCBI**, qui définit des identifiants commençant par préfixe “rs” suivi d’un nombre entier. Comme nous allons utiliser cette base de données pour la suite de ce TP, nous voudrions déjà connaître l’identifiant “rs” de ce variant fréquemment associé à la drépanocytose. On peut le retrouver sur le site OMIM.

Sur Ametice, répondez à la question suivante

- q.02omm.4: Quel est l’identifiant dbSNP (commençant par le préfixe rs) du variant génétique causal le plus fréquent pour la drépanocytose ?

Astuce : pour accéder à la description d’un variant et aux références vers d’autres bases de données, cliquez sur son identifiant OMIM

Cliquez ici

#603903
Table of Contents

Title
Phenotype–Gene Relationships
Clinical Synopsis
Text
Description
Clinical Features
Diagnosis
Clinical Management
Pathogenesis
Molecular Genetics
Population Genetics
Animal Model
History
References
Contributors
Creation Date
Edit History

▼ **Molecular Genetics**

The most common cause of sickle cell anemia is HbS (141900.0243), with hemoglobin SS disease being most prevalent in Africans. Rees et al. (2010) listed genotypes that had been reported to cause sickle cell disease. +

Modifier Genes

Priapism, a vasoocclusive manifestation of sickle cell disease, affects more than 30% of males with the disorder. In sickle cell anemia patients, 148 with priapism and 529 without, Nolan et al. (2005) searched SNPs from 44 genes of different functional classes for an association with priapism. By genotypic and haplotype analysis, they found an association between SNPs in the KLOTHO gene (604824) and priapism (rs2249358 and rs211239; adjusted odds ratio of 2.6 and 1.7, respectively). Nolan et al. (2005) noted that the finding may have broader implications in sickle cell disease, as the KL protein regulates vascular functions, including the expression of VEGF (192240) and release of endothelial nitric oxide (see 163729). +

Sickle cell anemia is phenotypically complex, with different clinical courses ranging from early childhood mortality to a virtually unrecognized condition. Overt stroke is a severe complication affecting 6 to 8% of individuals with sickle cell anemia. Modifier genes might interact to determine the susceptibility to stroke. Using Bayesian networks, Sebastiani et al. (2005) analyzed 108 SNPs in 39 candidate genes in 1,398 individuals with sickle cell anemia. They found that 31 SNPs in 12 genes interacted with fetal hemoglobin to modulate the risk of stroke. This network of interactions included 3 genes in the TGF-beta pathway (see 190180) and SELP (173610). Sebastiani et al. (2005) validated their model in a different population by predicting the occurrence of stroke in 114 individuals with 98.2% accuracy. +

Uda et al. (2008) found that the C allele of rs11886868 in the BCL11A gene (606557.0002) was associated with an ameliorated phenotype in patients with sickle cell anemia, due to increased production of fetal hemoglobin. +

3. UCSC genome browser : analyse génomique du gène de l'hémoglobine B et du variant rs334

Lorsqu'on examine les variations génétiques, il est important de les interpréter dans le contexte de leur environnement génomique, pour comprendre leurs conséquences phénotypiques. La structure des gènes eucaryotes inclut des introns et des exons. Chez les métazoaires et les plantes, les gènes sont séparés par de larges régions intergéniques. L'impact d'une variation dépend fortement de sa localisation intra- ou intergénique.

Le navigateur de génomes UCSC Genome Browser (genome-euro.ucsc.edu) est un outil en ligne qui permet de visualiser et d'explorer des données génomiques pour une grande variété d'organismes. Il offre une interface interactive pour accéder à des annotations, comparer des séquences et analyser des régions spécifiques du génome. Comme nous l'avons vu lors d'un TP précédent, UCSC genome browser comporte de nombreuses pistes d'annotation. Nous allons consulter celles consacrées aux variations génétiques.

3. UCSC genome browser : analyse génomique du gène de l'hémoglobine B et du variant rs334

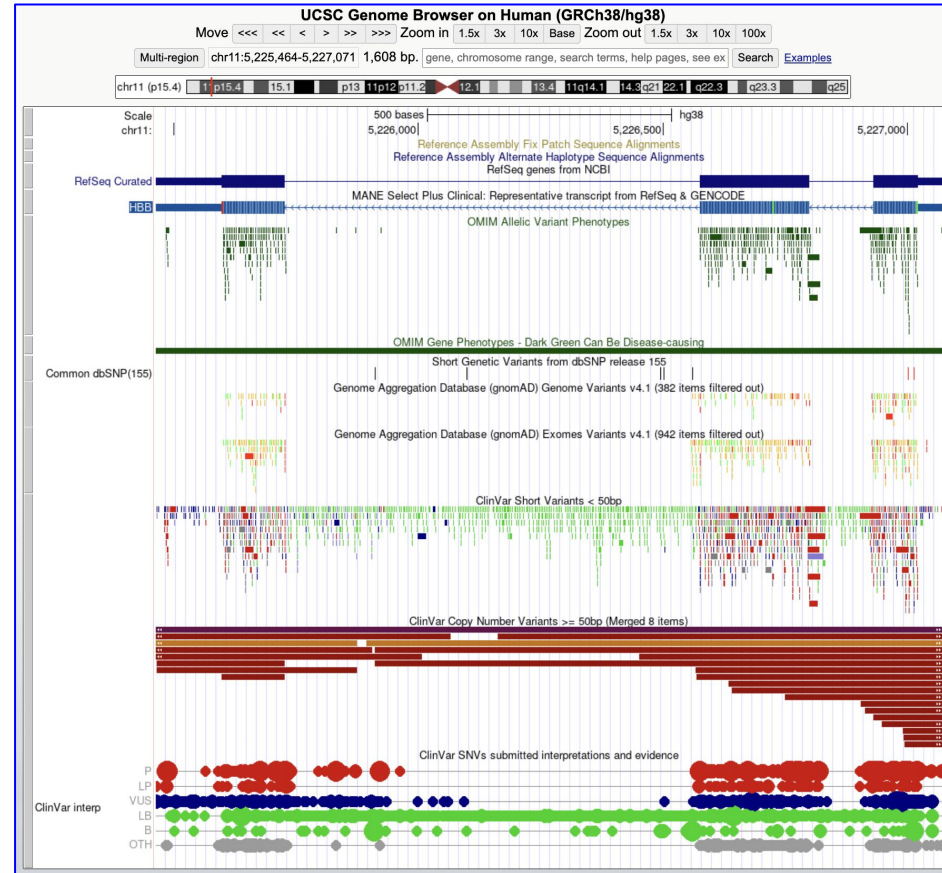
Le site OMIM nous a indiqué le chromosome et la bande cytogénétique du gène HBB. Le navigateur de génomes nous permet d'obtenir la position précise du gène et d'analyser sa structure.

- Ouvrez une connexion au **UCSC Genome Browser** (genome-euro.ucsc.edu).
- Sélectionnez la version hg38 du génome humain (**Genomes** → **Human GRCh38/hg38**).
- Effectuez une recherche pour localiser le gène **HBB**, en utilisant les annotations **MANE Select Plus Clinical**.

Astuce : UCSC offre une série de pistes d'annotation spécialisées pour les variations génomiques, et pour les phénotypes. Pour avoir une idée des informations disponibles concernant les variants, vous pouvez éventuellement activer certaines de ces pistes, par exemple OMIM, ClinVar, dbSNP, GnomAD.

Sur Ametice, répondez aux questions suivantes

10. q.03ucsc.1: Dans le UCSC Genome Browser, quelle est la position précise du gène humain HBB ?
11. q.03ucsc.2: Quel est le sens de transcription du gène HBB ?
12. q.03ucsc.3: En utilisant le UCSC Genome Browser, lequel des éléments suivants est situé en amont du gène *HBB* et joue un rôle crucial dans sa régulation ?
13. q.03ucsc.4: Combien de régions exoniques codantes peut-on observer dans le UCSC Genome Browser pour le gène HBB ?



3. UCSC genome browser : analyse génomique du gène de l'hémoglobine B et du variant rs334

Le navigateur des génomes UCSC permet aux utilisateurs de visualiser l'emplacement exact des variants dans le contexte des gènes, des régions régulatrices, et d'autres éléments fonctionnels, facilitant ainsi l'interprétation de leur impact biologique. Nous allons nous intéresser au variant rs334, responsable de la drépanocytose.

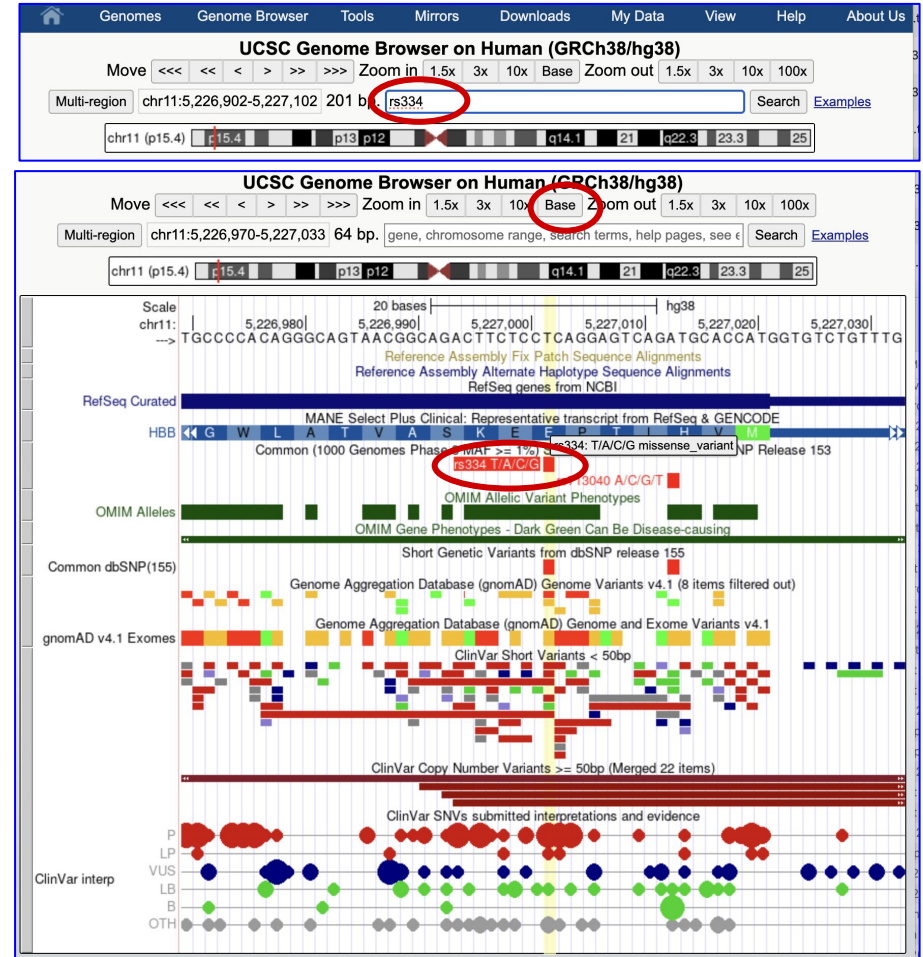
- Connectez-vous au **UCSC Genome Browser** (genome-euro.ucsc.edu)
- Sélectionnez la version hg38 du génome humain (Genomes → **Human GRCh38/hg38**)
- Recherchez le variant **rs334**
- Sélectionnez : **Common (1000 Genomes Phase 3 MAF >= 1%)**

Sur Amétice, répondez aux questions suivantes.

14. q.03ucsc.5 A quelle position du génome se trouve le variant rs334 ?
15. q.03ucsc.6 Où se situe le variant rs334 par rapport au gène HBB ?

Astuces :

- Vous pouvez localiser un variant en entrant son ID dbSNP dans la boîte de recherche
- Au-dessus de la carte génomique, vous pouvez cliquer sur "Base" pour zoomer au niveau du nucléotide.



4. dbSNP : fréquence populationnelle des variants

Un indice important de la pathogénicité des variants génétiques est leur fréquence dans les différentes populations humaines. Par exemple, on peut s'attendre à qu'un variant fréquent ne cause généralement pas de maladies, et à ce que les variants pathogéniques soient plutôt rares, puisqu'ils font généralement l'objet d'une contre-sélection. Par ailleurs, il est possible qu'un variant aie des fréquences distinctes dans des populations qui seraient soumises à des contraintes évolutives différentes du fait de leur environnement.

La base de données de référence des variants génétiques est **dbSNP**, gérée par le NCBI. dbSNP est une ressource de référence qui stocke des informations sur les variations génétiques, notamment les polymorphismes nucléotidiques (**single-nucleotide polymorphisms, SNP**) dans le génome humain et dans d'autres organismes. Elle fournit des données précieuses telles que les allèles alternatifs et leurs fréquences dans différentes populations. Nous allons la consulter pour analyser les fréquences populationnelles du variant rs334.

4. dbSNP : fréquence populationnelle des variants

Nous allons consulter dbSNP pour analyser les fréquences populationnelles du variant rs334. Le site dbSNP est organisé avec une boîte bleu qui résume quelques informations sur le variante et puis des onglets plus bas avec des informations détaillés: “Frequency”, et “Variant details”.

- Connectez-vous à la base de données **dbSNP** (www.ncbi.nlm.nih.gov/snp)
- Effectuez une recherche avancée de l’identifiant de SNP (**Reference SNP ID**) "**rs334**".
- **Cliquez sur l’ID rs334** de la première réponse (note: d’autres identifiants qui avaient été assignés à la même variation ont été fusionnés avec rs334)
- Explorez les onglets “**Frequency**” et “**Variant Details**” pour répondre aux questions.

Sur Ametice, répondez aux questions suivantes.

16. q.04snp.1 Quel est le nucléotide à cette position dans le génome de référence?
17. q.04snp.2 Quel est l'acide aminé à cette position dans la protéine HBB de référence?
18. q.04snp.3 Quelle est la position de cette AA dans la protéine HBB ?
19. q.04snp.4 Quels nucléotides alternatifs ont été observés à cette position ?
20. q.04snp.5 Quelle est la conséquence du variant génétique rs334 ?
21. q.04snp.6 Quelle est la fréquence de l'allèle A dans la population européenne ?
22. q.04snp.7 Quelle est la fréquence de l'allèle A dans la population africaine ?
23. q.04snp.8 Est-ce que la fréquence du variant rs334 chez les européens et africains pourrait expliquer la différence d’incidence de drépanocytose dans ces deux populations ?

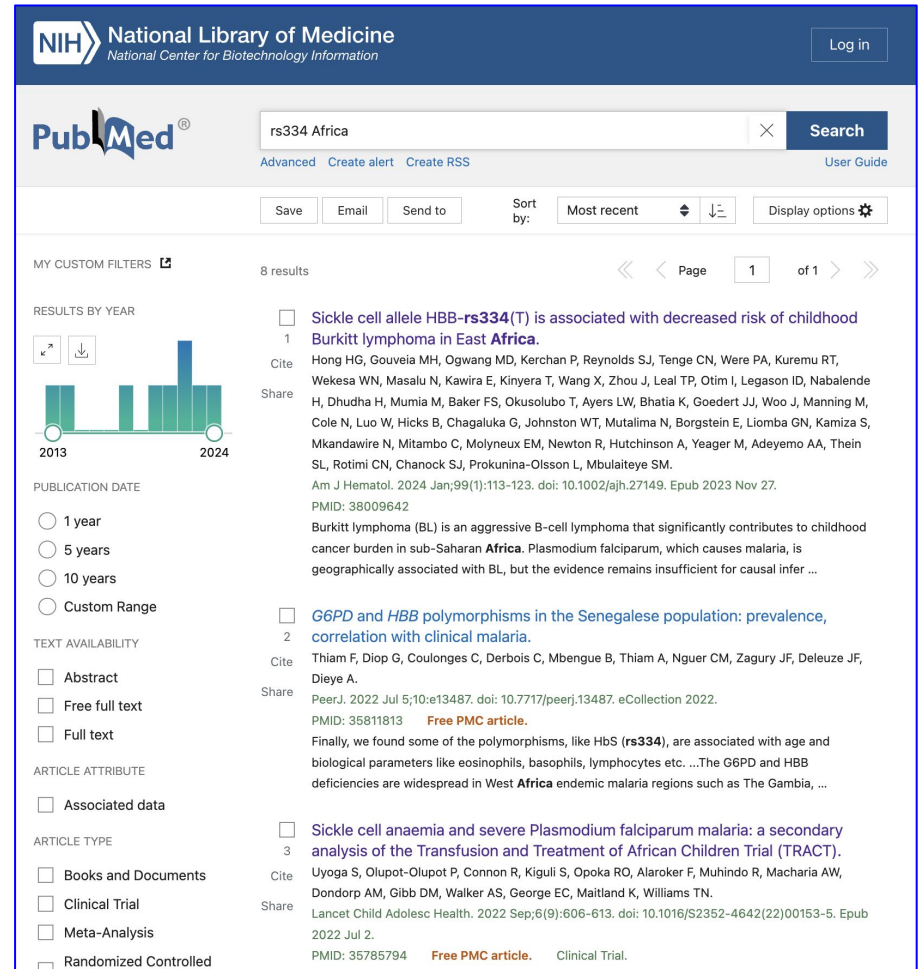
5. Pubmed: impact du variant rs334 dans les populations africaines

Pour comprendre la différence de fréquence entre populations africaines et autres populations, nous allons consulter la bibliographie sur Pubmed, la base de données spécialisée pour les publications scientifiques en biologie et santé.

- Connectez-vous à PubMed: (pubmed.ncbi.nlm.nih.gov)
- Saisissez dans le champ de recherche: rs334 Africa

Sur Ametice, répondez aux questions suivantes.

24. q.05pm.1 Sur la base des publications, quelles maladies sont influencées par le variant rs334 dans la population africaine ? (une ou plusieurs réponses)
25. q.05pm.2 A partir du résumé de la publication avec PMID 35785794, sélectionner les réponses vraies.



NIH National Library of Medicine
National Center for Biotechnology Information

rs334 Africa Search

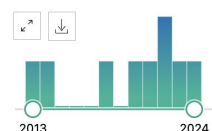
Advanced Create alert Create RSS User Guide

Save Email Send to Sort by: Most recent Display options

MY CUSTOM FILTERS

8 results

RESULTS BY YEAR



PUBLICATION DATE

1 year
 5 years
 10 years
 Custom Range

TEXT AVAILABILITY

Abstract
 Free full text
 Full text

ARTICLE ATTRIBUTE

Associated data

ARTICLE TYPE

Books and Documents
 Clinical Trial
 Meta-Analysis
 Randomized Controlled

1 Sickle cell allele HBB-rs334(T) is associated with decreased risk of childhood Burkitt lymphoma in East Africa.
Cite Hong HG, Gouveia MH, Ogwang MD, Kerchan P, Reynolds SJ, Tenge CN, Were PA, Kuremu RT, Wekesa WN, Masalu N, Kawira E, Kinyera T, Wang X, Zhou J, Leal TP, Otim I, Legason ID, Nabalende H, Dhudha H, Mumia M, Baker FS, Okusolubo T, Ayers LW, Bhatia K, Goedert JJ, Woo J, Manning M, Cole N, Luo W, Hicks B, Chagaluka G, Johnston WT, Mutalima N, Borgstein E, Liomba GN, Kamiza S, Mkandawire N, Mitambo C, Molyneux EM, Newton R, Hutchinson A, Yeager M, Adeyemo AA, Thain SL, Rotimi CN, Chanock SJ, Prokunina-Olsson L, Mbulaiteye SM.
Am J Hematol. 2024 Jan;99(1):113-123. doi: 10.1002/ajh.27149. Epub 2023 Nov 27.
PMID: 38009642
Burkitt lymphoma (BL) is an aggressive B-cell lymphoma that significantly contributes to childhood cancer burden in sub-Saharan Africa. Plasmodium falciparum, which causes malaria, is geographically associated with BL, but the evidence remains insufficient for causal infer ...

2 G6PD and HBB polymorphisms in the Senegalese population: prevalence, correlation with clinical malaria.
Cite Thiam F, Diop G, Coulonges C, Derbois C, Mbengue B, Thiam A, Nguer CM, Zagury JF, Deleuze JF, Dieye A.
Share PeerJ. 2022 Jul 5;10:e13487. doi: 10.7717/peerj.13487. eCollection 2022.
PMID: 35811813 Free PMC article.
Finally, we found some of the polymorphisms, like HbS (rs334), are associated with age and biological parameters like eosinophils, basophils, lymphocytes etc. ...The G6PD and HBB deficiencies are widespread in West Africa endemic malaria regions such as The Gambia, ...

3 Sickle cell anaemia and severe Plasmodium falciparum malaria: a secondary analysis of the Transfusion and Treatment of African Children Trial (TRACT).
Cite Uyoga S, Olupot-Olupot P, Cannon R, Kiguli S, Opoka RO, Alaroker F, Muhindo R, Macharia AW, Dondorp AM, Gibb DM, Walker AS, George EC, Maitland K, Williams TN.
Share Lancet Child Adolesc Health. 2022 Sep;6(9):606-613. doi: 10.1016/S2352-4642(22)00153-5. Epub 2022 Jul 2.
PMID: 35785794 Free PMC article. Clinical Trial.

Au cours de ce TP, nous nous sommes intéressés à une maladie monogénique (la drépanocytose). Nous avons consulté différents sites Web pour collecter des informations générales sur la maladie, telles que ses symptômes, sa prévalence dans différentes populations, et son mode de transmission héréditaire. En particulier, la base de données OMIM nous a permis de faire le lien entre la maladie, les gènes et les variants génétiques qui y contribuent.

Nous avons ensuite utilisé le navigateur de génomes UCSC pour explorer la structure des gènes impliqués, leur position dans le génome et la position des variants connus pour la maladie. Nous nous sommes intéressés à un variant génétique particulier (rs344) et avons recherché les allèles observés pour ce variant. Nous avons aussi examiné la fréquence des allèles de ce variant dans différentes populations.

Enfin, nous avons effectué une recherche dans la littérature scientifique (PubMed) pour essayer de comprendre pourquoi ces variants impliqués dans la maladie présentent de fortes différences de fréquences entre différentes populations. Ceci nous a amenés à constater qu'un variant génétique peut dans certains cas combiner des effets délétères et protecteurs pour les individus qui le portent. Ces effets peuvent varier en fonction des populations, car celles-ci sont soumises à des contraintes évolutives différentes en fonction de leur environnement.

Au fil de ces étapes, vous devriez avoir acquis la compétence de consulter une série de bases de données et de connaissances pour obtenir une vision intégrée de la relation entre séquence génomique, variations génétiques, conséquence moléculaire d'une mutation et impact sur la santé. Cette compétence pourra vous servir pour analyser les composantes génétiques d'autres pathologies humaines.