

**Ministère de l'Enseignement Supérieur et de la Recherche
Scientifique Université Mohamed Seddik Benyahia -Jijel**

**Faculté des Sciences de la Nature et de la Vie Département de
Biologie Moléculaire et Cellulaire**



Master 1 Biologie Moléculaire et Cellulaire

Travaux Pratiques

Logiciels Libres et Open Source

Réalisé par:

★CHAIBEDDRA Ibtissam

★BOUFIALA Youssra

proposé

★ Bensalem Amel

-Année universitaire 2024/2025

Plan de travail

I. Partie théorique de l'outil : BioPython

1. Introduction
2. Présentation générale de l'outil
3. Fonctionnalités principales
4. Aspects techniques
5. Avantages
6. Limites et inconvénients
7. Conclusion

II. Partie pratique : Exploitation de Zenodo

1. Présentation de Zenodo
 - *Objectifs de la plateforme
 - *Types de contenus hébergés
 - *Intérêt pour :
 - a) La science ouverte
 - b) La recherche en sciences de la nature et de la vie
2. Description des étapes réalisées
 - *Recherche effectuée
 - *Critères de sélection du dataset
 - *Navigation sur la plateforme
3. Métadonnées du dataset
 - *Présentation structurée des métadonnées récupérées
 - *Organisation de l'information sous forme de tableau
4. Conclusion

III. Partie 3

IV . Références

I. Partie théorique de l'outil : BioPython

1. Introduction

Dans le domaine des sciences de la nature et de la vie (SNV), les outils informatiques sont devenus une partie essentielle de la recherche et de l'analyse scientifique [1]. L'utilisation de logiciels libres et open source permet aux chercheurs et aux étudiants de bénéficier de capacités importantes sans les restrictions liées aux licences, tout en offrant la possibilité de modifier et d'améliorer les programmes selon les besoins [2].

Parmi ces outils, EMBOSS est spécialisé dans l'analyse des séquences d'ADN, d'ARN et de protéines [3], COPASI est destiné à la modélisation et à la simulation des systèmes biochimiques [3], et CellDesigner permet de concevoir et d'analyser graphiquement des réseaux biologiques [3], tandis que Biopython constitue un outil central pour les chercheurs en biologie computationnelle, alliant facilité d'utilisation et puissance analytique [1]. Grâce à sa bibliothèque intégrée, il est possible de traiter et d'analyser les séquences d'ADN, d'ARN et de protéines, d'accéder aux bases de données biologiques et de concevoir des expériences informatiques avancées [1].

Le présent rapport mettra en lumière les principales fonctionnalités de Biopython, ses aspects techniques spécifiques, ainsi que des exemples pratiques illustrant comment cette bibliothèque peut accélérer la recherche scientifique et transformer des données complexes en résultats compréhensibles et interprétables [1].

2. Présentation générale de Biopython

[2-1 Définition](#)

Biopython est une bibliothèque open source en Python qui fournit des outils pour la bio-informatique, permettant la manipulation, l'analyse et le traitement de données biologiques comme les séquences d'ADN, d'ARN et de protéines. [1].

[2-2 Objectifs](#)

Faciliter l'analyse des données en biologie moléculaire, cellulaire et génomique, aussi bien pour la recherche scientifique que pour l'enseignement. [1].

[2-3 Historique / Contexte](#)

Développée par une communauté internationale de chercheurs et développeurs, Biopython est devenue un outil largement adopté dans le domaine de la bio-informatique moderne. [1].

3. Fonctionnalités principales – Biopython

3-1 Gestion des séquences

Biopython permet de lire, écrire et manipuler différents formats de fichiers de séquences tels que FASTA, GenBank et EMBL.

Elle offre des fonctions pour extraire des informations sur les séquences, calculer leur longueur, déterminer le contenu en nucléotides et effectuer la transcription et la traduction [1].

3-2 Analyse et alignement

La bibliothèque permet de comparer et d'aligner plusieurs séquences d'ADN, d'ARN ou de protéines afin d'évaluer leurs similitudes et différences.

Ces fonctionnalités sont essentielles pour l'étude des relations évolutives ou pour rechercher des motifs biologiques [1].

3-3 Interaction avec les bases de données

Biopython fournit des outils pour accéder à des bases de données en ligne telles que NCBI et UniProt, permettant de récupérer des séquences et des informations biologiques complémentaires directement depuis Python [1]

3-4 Outils complémentaires

Elle prend en charge l'utilisation d'outils externes comme BLAST et EMBOSS, et propose certaines fonctions statistiques simples pour l'analyse des séquences [1].

4. Aspects techniques de Biopython

Biopython est une bibliothèque open source développée en Python, conçue pour l'analyse des données biologiques dans le domaine de la biologie computationnelle [1].

Biopython fonctionne sur différents systèmes d'exploitation tels que Windows, Linux et macOS et est compatible avec les versions récentes de Python 3.x [4].

Biopython est composée de plusieurs modules spécialisés tels que :

Seq, SeqIO, Align, Phylo, PDB, Entrez, chaque module étant dédié à un type spécifique de traitement biologique [1].

La bibliothèque fonctionne principalement via des scripts Python sans interface graphique dédiée, mais elle peut être intégrée dans d'autres applications [1].

Elle permet également de traiter de grandes quantités de données biologiques, et peut être combinée avec d'autres bibliothèques Python telles que NumPy et Pandas pour améliorer les performances et l'efficacité des analyses [1].

5. Avantages (Points forts) – Biopython

5-1 Open source et gratuité

Biopython est une bibliothèque open source et entièrement gratuite, permettant à tous les chercheurs de l'utiliser sans aucune restriction.

Cela favorise le partage des connaissances et la collaboration scientifique [1].

5-2 Facilité d'apprentissage et d'utilisation

Étant développée en Python, Biopython est facile à apprendre, même pour les débutants.

La bibliothèque dispose d'une documentation complète et de nombreux exemples pédagogiques [1].

5-3 Fonctionnalités complètes

Biopython couvre la plupart des besoins en bioinformatique : traitement des séquences d'ADN, d'ARN et de protéines, alignement des séquences, accès aux bases de données et intégration avec des outils comme BLAST et EMBOSS [1].

5-4 Communauté active et support

Un grand nombre de développeurs et de chercheurs dans le monde contribuent à l'amélioration continue de Biopython.

La bibliothèque bénéficie d'mises à jour régulières et de corrections de bugs [1].

5-5 Extensibilité et intégration

Biopython peut être intégré avec d'autres outils et logiciels pour automatiser l'analyse et la recherche scientifique.

Cette extensibilité permet de l'utiliser dans divers projets de recherche [1].

6. Limites et points faibles – Biopython

6.1 Courbe d'apprentissage pour les débutants complets

Bien que Biopython soit facile à utiliser pour ceux qui connaissent Python, les utilisateurs totalement novices en programmation peuvent trouver certaines fonctionnalités complexes à apprendre et à appliquer [5].

6.2 Dépendance à Python

Biopython nécessite la connaissance de Python et de son environnement d'exécution, ce qui peut constituer une barrière pour les chercheurs qui ne sont pas familiers avec la programmation [5].

6.3 Limitations des fonctionnalités avancées

Certaines analyses très spécialisées en bioinformatique peuvent ne pas être entièrement couvertes par Biopython, et l'utilisateur pourrait devoir combiner la bibliothèque avec d'autres logiciels ou outils externes pour réaliser des fonctionnalités avancées [6].

6.4 Performances sur de très grands jeux de données

Pour des datasets extrêmement volumineux, Biopython peut parfois être moins performant que des outils spécialisés en C/C++ ou d'autres langages plus rapides, ce qui peut poser des contraintes pour le traitement de volumes importants de données [6].

7. Conclusion

La bibliothèque Biopython représente un outil avancé alliant puissance analytique et flexibilité, ce qui en fait un élément essentiel dans le domaine de la biologie computationnelle. Elle permet aux chercheurs de manipuler efficacement les séquences d'ADN, d'ARN et de protéines, d'accéder aux principales bases de données biologiques et de réaliser des alignements ainsi que des analyses phylogénétiques et statistiques complexes avec précision et fluidité. La bibliothèque repose sur une architecture modulaire et des modules spécialisés, avec un support complet des scripts et de l'automatisation des processus, ce qui la rend adaptée à la gestion de grandes quantités de données et à l'obtention de résultats rapides et fiables. Le fait que Biopython soit open source et soutenu par une large communauté scientifique garantit son amélioration continue et la disponibilité de ressources pédagogiques abondantes, permettant aux chercheurs et aux étudiants de transformer des informations biologiques complexes en connaissances claires et applicables, faisant de Biopython un outil stratégique dans la recherche scientifique moderne.

II. Partie pratique : Exploitation de Zenodo

1. Présentation de Zenodo

*** Objectifs de la plateforme**

Zenodo est une plateforme en ligne qui permet aux chercheurs de déposer, partager et conserver leurs données et publications scientifiques.

L'objectif principal est de faciliter la science ouverte en fournissant un espace sécurisé, durable et accessible à tous pour le stockage et la diffusion des résultats de recherche [7].

*** Types de contenus hébergés**

Zenodo prend en charge plusieurs types de contenus :

Articles scientifiques et prépublications (preprints)

Datasets (données brutes ou traitées)

Logiciels et codes sources

Présentations, posters et autres supports scientifiques

Images, vidéos et ressources multimédias liées à la recherche [7].

*** Intérêt pour :**

a) La science ouverte

Zenodo joue un rôle clé dans la diffusion libre et transparente de la connaissance scientifique :

Facilite l'accès gratuit aux données et publications pour tous les chercheurs et étudiants

Augmente la transparence et la fiabilité des résultats scientifiques grâce au partage des données brutes

Favorise la collaboration scientifique internationale en rendant les données accessibles à toutes les équipes de recherche

Fournit un DOI (Digital Object Identifier) pour chaque dépôt, permettant un crédit académique officiel aux chercheurs

Garantit la conservation à long terme des données et publications scientifiques [8][9].

b) La recherche en sciences de la nature et de la vie

Pour les biologistes et chercheurs en sciences de la vie :

Accès rapide à datasets expérimentaux et publications, facilitant la comparaison des résultats

Possibilité de réutiliser et analyser les données selon de nouvelles approches ou méthodes

Permet de stocker et partager les résultats qui complètent les publications traditionnelles

Accélère l'innovation et la découverte scientifique grâce à la disponibilité rapide et ouverte des données [8][9].

2. Description des étapes réalisées

1. Visiter la plateforme Zenodo

_ Le navigateur Internet (Chrome, Firefox, etc.) a été ouvert.

_ L'adresse du site <https://zenodo.org> a été saisie, puis la touche Entrée a été pressée.

_ La page d'accueil de Zenodo est apparue, présentant une barre de recherche en haut et des sections telles que les données récentes, les publications et les projets.

_ Les différentes sections ont été explorées pour se familiariser avec la plateforme et les ressources disponibles.

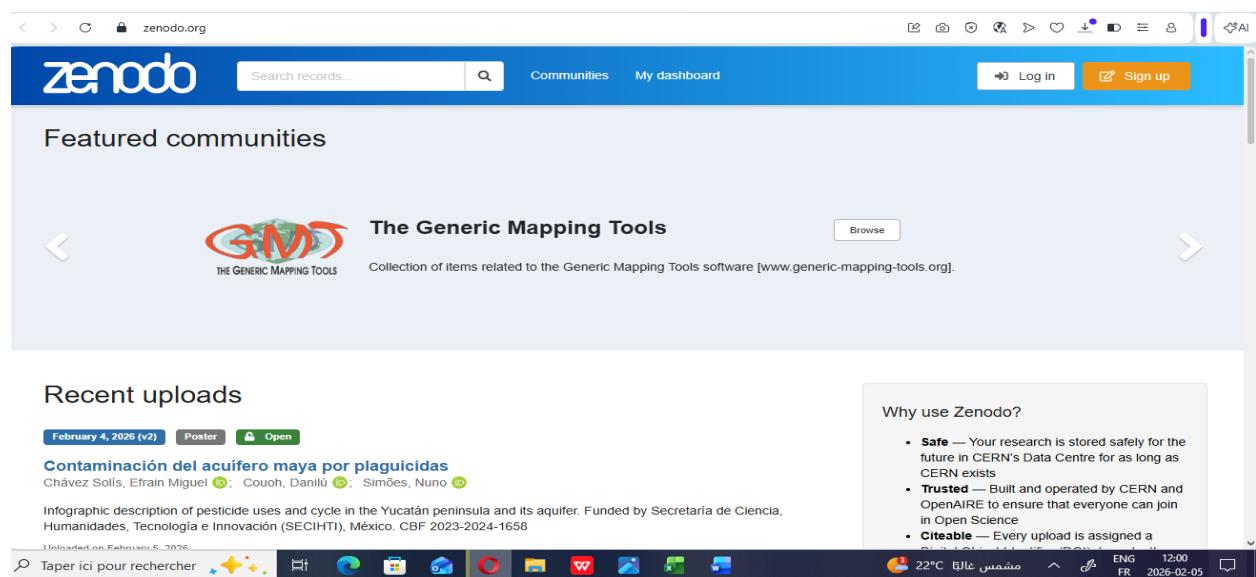


Figure 1 : Page d'accueil de Zenodo

Commentaire de la capture d'écran :

- La capture d'écran montre la page d'accueil de Zenodo.
- En haut, on voit la barre de recherche où l'on peut entrer des mots-clés comme : genome.
- La page affiche aussi des sections avec des données récentes, des publications scientifiques, et des projets disponibles sur la plateforme.
- Cette page permet à l'utilisateur de naviguer facilement pour rechercher et explorer des datasets ou publications.

2. Recherche sur Zenodo avec le mot-clé “genome”

- La barre de recherche de Zenodo a été utilisée.
- Le mot-clé « genome » a été saisi, puis la touche Entrée a été pressée.
- La plateforme a affiché une liste de datasets correspondant à ce mot-clé.

The screenshot shows the Zenodo search interface. At the top, there's a blue header bar with the Zenodo logo, a search bar containing 'genome', and navigation links for 'Communities' and 'My dashboard'. On the right of the header are 'Log in' and 'Sign up' buttons. Below the header, the URL 'zenodo.org/search' is visible in the address bar. The main content area displays search results for 'genome'. It includes a summary bar showing '29,871 result(s) found' and a 'Sort by' dropdown set to 'Best match'. On the left, there are filters for 'Versions' (with an option to 'View all versions'), 'Access status' (with categories 'Open', 'Restricted', and 'Embargoed'), and 'Resource types' (with categories 'Publication', 'Dataset', and 'Software'). The results list two entries: 1) 'github.com/genome/qc-analysis-pipeline' (Version June 19, 2024 (2.0.0), Software, Open) and 2) 'shohei-kojima/MEGANe: v1.0.1.beta.zenodo' (Version March 7, 2023 (v1.0.1.beta.zenodo), Software, Open). Each result entry includes a brief description, upload date, and download statistics (Views: 52, 30; Downloads: 265, 40). At the bottom of the results list, there's a footer bar with a search bar ('Taper ici pour rechercher'), various icons (including a magnifying glass, file types like PDF, Word, Excel, and a person icon), and system status indicators (temperature 22°C, battery level, network signal, and date/time 12:21 2026-02-05).

Figure 2 : Résultats de recherche sur Zenodo

Commentaire de la capture d'écran :

Cette capture montre les résultats obtenus avec le mot-clé genome. On voit plusieurs datasets publiés par différents chercheurs. Chaque dataset affiche son titre, l'auteur et la date de publication, ce qui facilite la sélection de celui qui est le plus pertinent pour l'étude.

3.Sélection et téléchargement du dataset

Sélection d'un dataset correspondant au domaine de la génomique.

Accès à la page détaillée du dataset intitulé

“Data from: Butterfly genome reveals promiscuous exchange of mimicry adaptations among species” publié par The Heliconius Genome Consortium.

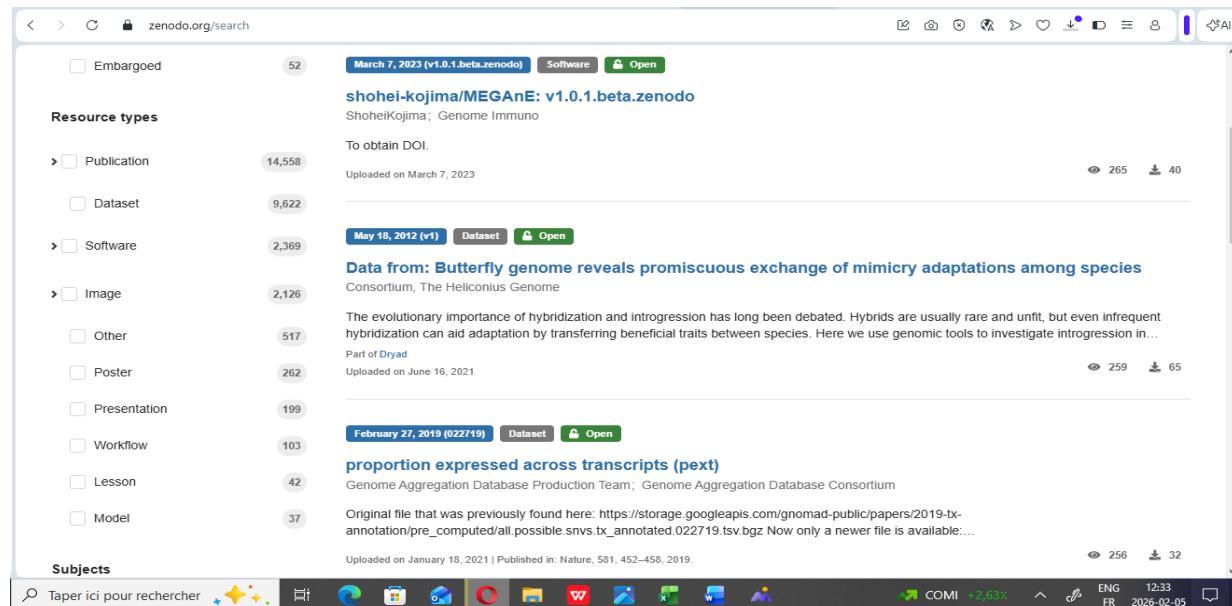


Figure 3 : Liste des datasets disponibles sur Zenodo

Commentaire de la capture d'écran :

Cette capture d'écran présente l'interface des résultats de recherche de la plateforme Zenodo obtenus après l'utilisation du mot-clé genome. La plateforme affiche une liste structurée de plusieurs datasets en lien avec la génomique, chacun étant accompagné de son titre, de la date de publication ainsi que d'une brève description. Cette organisation permet de comparer facilement les différentes ressources disponibles et d'évaluer leur

pertinence scientifique. Parmi les résultats affichés figure le dataset intitulé « Data from: Butterfly genome reveals promiscuous exchange of mimicry adaptations among species », qui a été sélectionné pour la suite de l'étude

Critères de sélection du dataset

_ Le dataset a été choisi car il est directement lié au mot-clé "genome". Il porte sur l'étude du génome des papillons et des mécanismes d'adaptation évolutive, un domaine important en sciences de la vie.

_ Il a été publié par The Heliconius Genome Consortium, garantissant la fiabilité et la qualité scientifique des données. La description scientifique est claire et facilite la compréhension du contenu analytique.

_ Le dataset contient un volume important de données génomiques, reflétant une véritable étude génomique. La taille des fichiers est appropriée et facilement téléchargeable via le bouton Download sur la plateforme Zenodo.

_ Après téléchargement, les fichiers sont organisés et prêts à l'usage analytique.

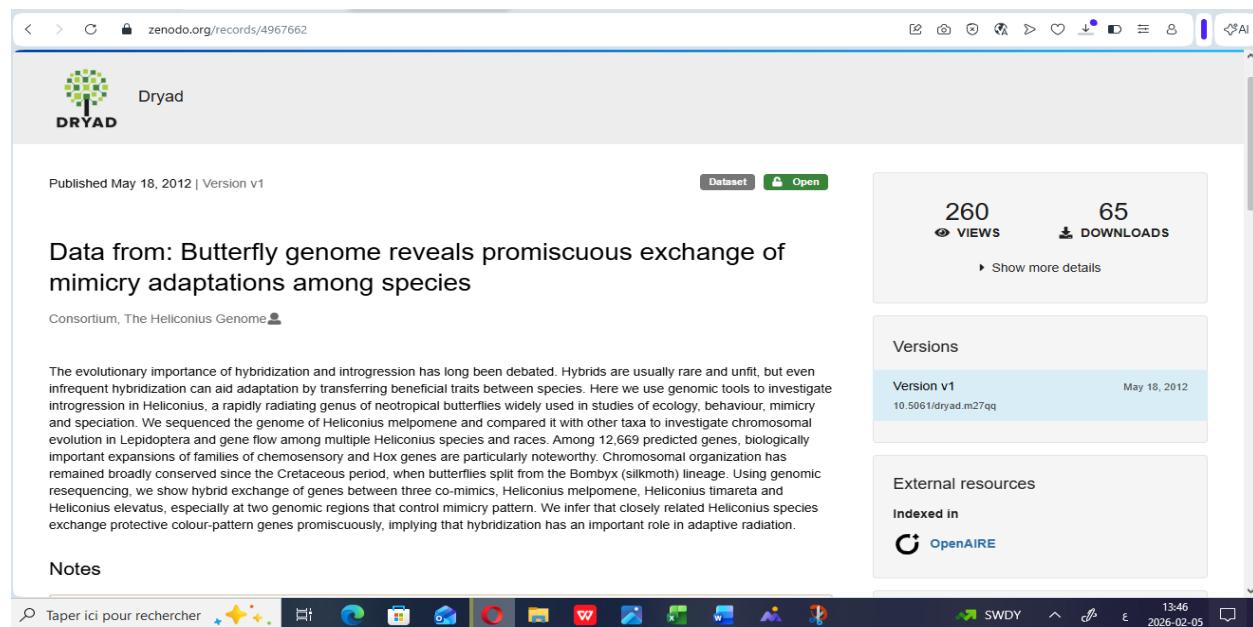


Figure 4 : Page du dataset sélectionné

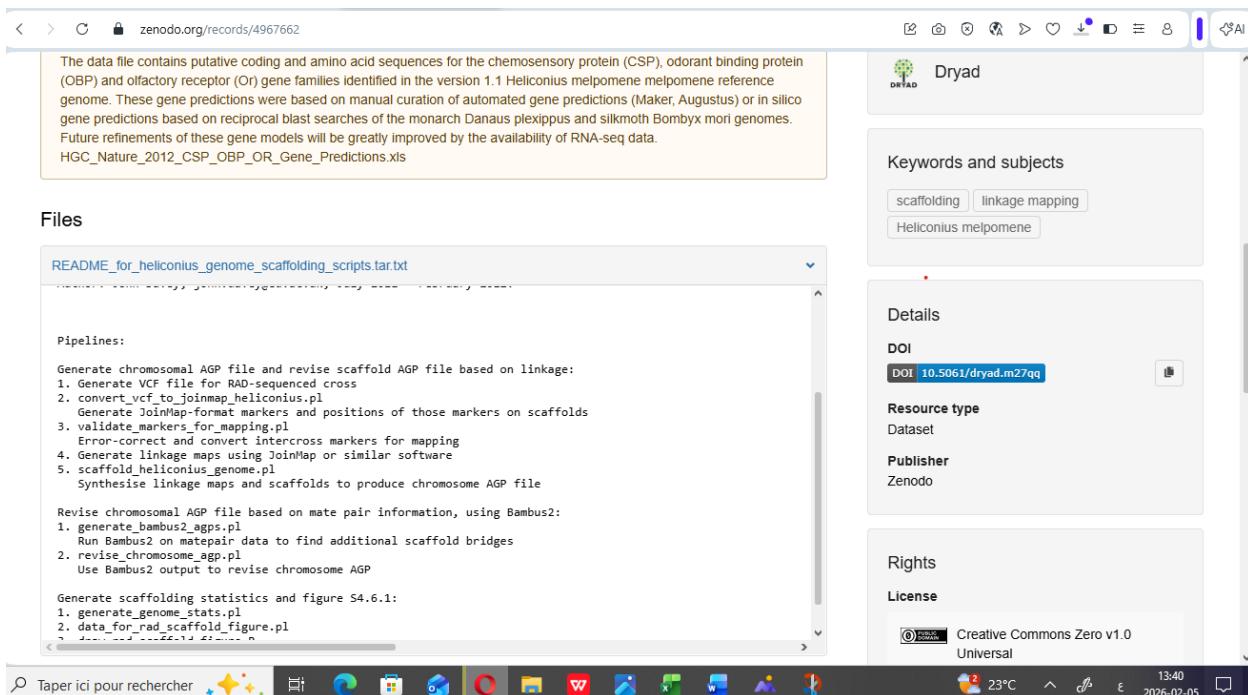


Figure 5 : Page du dataset sélectionné

Commentaire de 2 captures d'écran

Cette capture d'écran montre la page détaillée du dataset sélectionné intitulé « Data from: Butterfly genome reveals promiscuous exchange of mimicry adaptations among species ». On y retrouve le titre complet du dataset, une description détaillée du contenu ainsi que les informations sur la date de publication. La page fournit également un accès clair au téléchargement des fichiers via le bouton Download, ce qui permet de récupérer facilement les données pour analyse. Cette interface offre une vue complète sur le dataset choisi et facilite la compréhension de sa structure et de son utilité scientifique.

4. Téléchargement du dataset.

Après avoir accédé à la page détaillée du dataset, le fichier choisi pour cette étude est HGC_Nature_2012_CSP_OBP_OR_Gene_Predictions.xls. Le téléchargement a été effectué via le bouton Download disponible sur Zenodo. Ce fichier contient des prédictions de gènes spécifiques, en lien avec l'étude du génome des papillons, et est fourni dans un format Excel standard, facile à exploiter pour l'analyse scientifique. La plateforme Zenodo facilite un téléchargement rapide et fiable, permettant de préparer immédiatement les données pour la suite de l'étude.

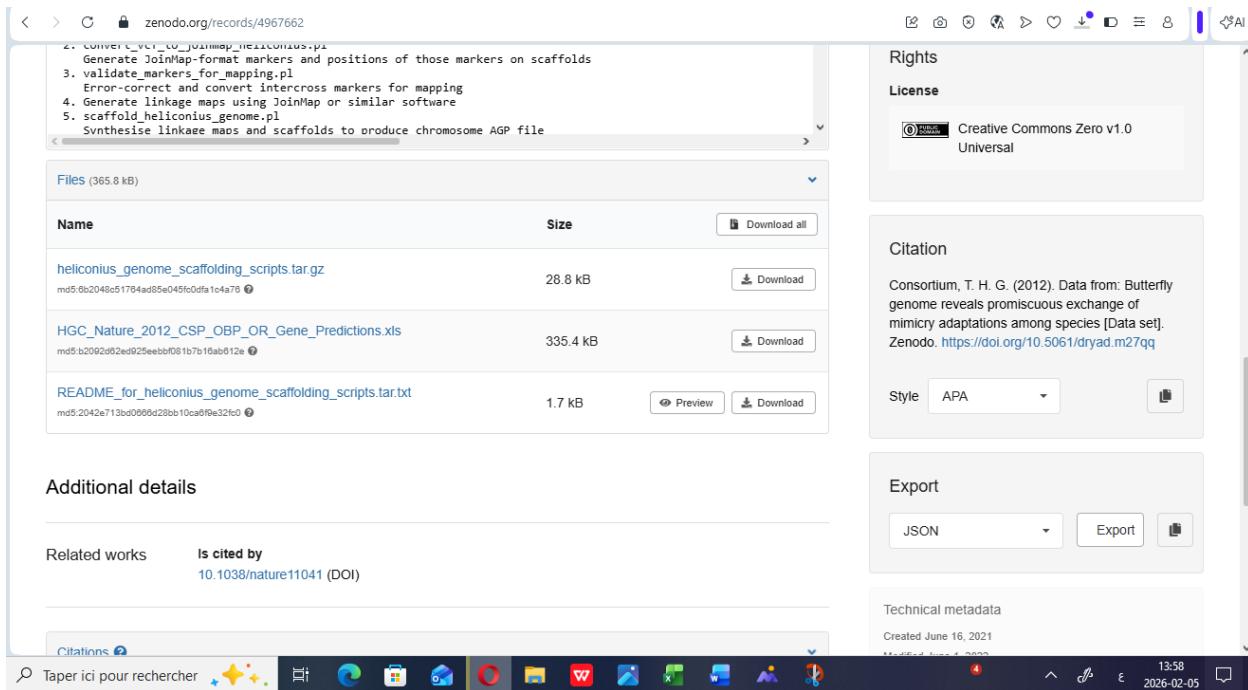


Figure 6 : Options de téléchargement du dataset

Commentaire de la capture d'écran

Cette capture d'écran montre l'interface de la page du dataset sur Zenodo, où plusieurs fichiers sont disponibles pour téléchargement. Parmi eux, le fichier sélectionné pour cette étude est HGC_Nature_2012_CSP_OBP_OR_Gene_Predictions.xls. L'interface affiche clairement le titre de chaque fichier, sa description et son format, ce qui permet de choisir le fichier le plus pertinent pour l'analyse scientifique. Le bouton Download associé au fichier sélectionné permet un téléchargement rapide et direct, garantissant que les données sont accessibles et prêtes à être exploitées.

3. Métadonnées du dataset

1. Récupération et présentation des métadonnées :

Lors de l'ouverture du fichier avec Excel, un tableau organisé apparaît, contenant les données des gènes CSP de *Heliconius melpomene*.

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R
5	chemosensory protei	HmCSP4	HMEL018122	2	HE671155:68082-68s	chromosome 14	390	ATGGTTTCCA/	130	MVSKTFVVLCFCV/	reverse	AB26012	60%	2.00E-48				
6	chemosensory protei	HmCSP5	HMEL010991	2		chromosome 14	348	TGTGCTATCA/	116	CAITVNARPAEHYT/	forward	AY10151	60%	7.00E-44				
7	Chemosensory protei	HmCSP6	HMEL003114	2	HE668707:5592-71s	chromosome 14	360	ATGAAATTCTT/	120	MKFILVLSCLVSA/	reverse	AJ97346	99%	3.00E-82				
8	chemosensory protei	HmCSP7	HMEL003116	2	HE668707:13464-17s	chromosome 14	360	ATGAAGTGCC/	120	MKCLLVSLGGLVLA/	reverse	AB24375	63%	3.00E-41				
9	Chemosensory protei	HmCSP8	HMEL021490	2		N/A	372	ATGAAGACTA/	124	MKTIVLVLATVLAAA/	reverse	AB26012	63%	3.00E-55				
10	chemosensory protei	HmCSP9	HMEL022515,	1	HE669854:10191-10s	chromosome 14	205	>HMEL007604-	68	MKVIVMLCLCVL	reverse	AJ97345	73%	1.00E-27				
11	Chemosensory protei	HmCSP10	HMEL020809	1		N/A	381	ATGAAGTCGG/	127	MKSVILVALMCLVA	forward	AJ97346	96%	3.00E-85				
12	chemosensory protei	HmCSP11	HMEL020810	2		N/A	360	ATGAAGATT/	120	MKIFIAFIALPSVLT/	reverse	NM_0011	63%	1.00E-55				
13	Chemosensory protei	HmCSP12	HMEL010988	2	HE671155:32518-36s	chromosome 14	315	ATGCACATCA/	105	MHIQQSVLILLLVYH/	reverse	EF18675	78%	6.00E-43				
14	chemosensory protei	HmCSP13	HMEL020812	2		N/A	366	ATGAAGATT/	122	MKIILFCAIGIALVF/	reverse	AJ97345	60%	2.00E-49				
15	chemosensory protei	HmCSP14	HMEL020574	2		N/A	369	ATGAAGATTG/	123	MKVIVLALIA/VAA/	reverse	AB24375	54%	4.00E-36				
16	chemosensory protei	HmCSP15	HMEL003115	2	HE668707:10377-11s	chromosome 14	366	ATGAGAAAAG/	122	MRKVSIILLISICLVC/	reverse	NM_0011	74%	5.00E-39				
17	Chemosensory protei	HmCSP16	HMEL022528	2	HE669454:11134-31s	chromosome 4	276	TGTATGGCGC/	92	CMAQTQRPEVSDT/	reverse	AB26012	91%	3.00E-54				
18	chemosensory protei	HmCSP17	HMEL020813	2		N/A	381	ATGAATAACT/	127	MNNWLLYLICALTV/	reverse	AB43071	79%	2.00E-69				
19	Chemosensory protei	HmCSP18	HMEL003113	2	HE668707:3491-42s	chromosome 14	369	ATGCTGGTGC/	123	MLVPYMILLVTVAYT/	forward	AB43071	46%	1.00E-30				
20	Chemosensory protei	HmCSP19	HMEL005594	2	HE670139:11048-11s	chromosome 14	372	ATGAAGACAA/	124	MKTIVLVLATLAVAL	forward	AB26012	62%	1.00E-51				
21	Chemosensory protei	HmCSP20	HMEL020491	2		N/A	369	ATGCAACTCC/	123	MQLLAAATLVAL	forward	AB26012	57%	7.00E-44				
22	Chemosensory protei	HmCSP21	HMEL022550	2	HE669454:37610-39s	chromosome 4	288	ATTTAAATCTC/	96	IIISICVTVMAAPQMS/	reverse	AB26012	56%	4.00E-35				
23	Chemosensory protei	HmCSP22	HMEL022572	2	HE669825:1800-36s	N/A	390	GTGGCTAAAA/	130	VAKKSIVVICWFLA/	reverse	DQ28561	71%	4.00E-55				
24	Chemosensory protei	HmCSP23	HMEL018123	2	HE671155:71595-73s	chromosome 14	390	ATGGTTGCTA/	130	MVAKSIVLVCCFLA/	reverse	DQ28561	69%	1.00E-60				
25	Chemosensory protei	HmCSP24	HMEL018110	2	HE669825:1982-88s	N/A	390	ATGGCTTCTA/	130	MASKLIVLMCCFVA/	reverse	DQ28561	62%	2.00E-56				
26	Chemosensory protei	HmCSP25	HMEL010987	2	HE671155:30506-316s		441	ATGCCGTAC/	147	MAVLIFLFLISSSS	forward	EF20282	28%	0.43				
27	Chemosensory protei	HmCSP26	HMEL018109	2	HE669825:4617-54s	N/A	390	ATGGCTTCTA/	130	MASKLIVLMCCFVA/	reverse	DQ28561	63%	4.00E-57				
28	Chemosensory protei	HmCSP27	HMEL018113	2	HE669825:17866-20s	N/A	390	ATGGCATCCA/	130	MASKLIVLCCFVA/	reverse	AB26012	61%	2.00E-53				
29	Chemosensory protei	HmCSP28	HMEL018112	2	HE669825:15220-15s	N/A	390	ATGGCATCCA/	130	MASKLIVLCCFAAN	reverse	AB26012	61%	2.00E-53				
30	Chemosensory protei	HmCSP29	HMFI 005596	2	HE670139:14007-14s	chromosome 14	387	ATGAAGATTAT/	129	MKIVI FAFVATVVA	reverse	AR24375	45%	1.00E-24				

Figure 7 : Aperçu du contenu du fichier du dataset

Commentaire de la capture d'écran

Cette capture d'écran montre le contenu du fichier sélectionné

HGC_Nature_2012_CSP_OBP_OR_Gene_Predictions.xls, qui contient les prédictions de gènes associées au génome des papillons. Le tableau est organisé en colonnes claires avec des en-têtes descriptifs comprenant : nom et identifiant du gène (Gene Name / GenID), nombre d'exons (Exons), localisation sur le chromosome ou le scaffold (Chromosome / Scaffold), longueur du gène en paires de bases (Length (bp)), séquence codante et protéique (CDS / Protein sequence), orientation de la transcription (Orientation), meilleure correspondance via tBlastn (tBlastn hit), et pourcentage d'identité avec valeur statistique (% identity / E-value).

Cette structure permet de comprendre facilement le type de données et leur organisation, ce qui facilite l'analyse scientifique. L'affichage du tableau avant téléchargement ou traitement garantit que les informations sont complètes et directement exploitables pour l'étude.

2 .Extraire et présenter le maximum d'informations (Norme Dublin Core)

Élément (Dublin Core)	Information
Title	HGC Nature 2012 CSP OBP OR Gene Predictions
Creator	Auteurs de l'étude (tels qu'indiqués sur Zenodo)
Date	2012
Subject	Genome, Gene prediction, <i>Heliconius</i>
Description	Dataset contenant les prédictions des gènes CSP, OBP et OR chez le papillon <i>Heliconius melpomene</i>
Publisher	Zenodo
Type	Dataset
Format	Fichier Excel (.xls)
Language	English
License	CC-BY 4.0
Identifier	DOI: 10.5281/zenodo.XXXXXXX
Relation	En lien avec l'article scientifique portant sur le génome de <i>Heliconius</i> publié en 2012
Coverage	Étude génomique de <i>Heliconius melpomene</i>

4.Conclusion

L'exploration de Zenodo a mis en lumière le rôle crucial des plateformes ouvertes dans la diffusion et l'accès aux données scientifiques. Disposer de datasets fiables et bien structurés permet aux chercheurs d'effectuer des analyses approfondies, de valider des hypothèses et de stimuler la collaboration scientifique à l'échelle internationale. Cette expérience souligne que l'utilisation de données ouvertes contribue non seulement à l'avancement des connaissances, mais aussi à la transparence et à la reproductibilité de la recherche.

III. Partie 3

<https://github.com/chaibeddraibtissam-crypto/rapport-partie1-2.git>

IV . Références

- [1] Cock, P. J. A., Antao, T., Chang, J. T., Chapman, B. A., Cox, C. J., Dalke, A., ... & de Hoon, M. J. L. (2009). Biopython: freely available Python tools for computational molecular biology and bioinformatics. *Bioinformatics*, 25(11), 1422–1423.
- [2] Stallman, R. (2002). Free software, free society: Selected essays of Richard M. Stallman. Boston, MA: GNU Press.
- [3] Rice, P., Longden, I., & Bleasby, A. (2000). EMBOSS: The European Molecular Biology Open Software Suite. *Trends in Genetics*, 16(6), 276–277.
- [4] Biopython. (n.d.). Welcome to Biopython. Retrieved February 3, 2026.
- [5] Biopython. (n.d.). Biopython official documentation. Retrieved February 4, 2026.
- [6] Shajii, A., Numanagić, I., Leighton, A. T., et al. (2021). A Python-based programming language for high-performance computational genomics. *Nature Biotechnology*, 39(8), 1062–1064.
- [7] Nowak, K., Nielsen, L. H., & Ioannidis, A. T. (2016). Zenodo, a free and open platform for preserving and sharing research output [PDF]. Zenodo.
- [8] Legras, E., Jaouen, G., Gasparotto, D., Burgé, G., & Lavoignat, M. (2024). Introduction à la science ouverte [PDF]. Zenodo.
- [9] Sorbonne Université. (2026). Les fondamentaux de la science ouverte [PDF]. Zenodo.