

Fonctionnement du site

Introduction : Vue d'ensemble et Objectifs

Ce projet consiste en la création d'un outil web permettant à la fois d'annoter des génomes bactériens et d'en faire une analyse fonctionnelle.

Notre outil est constitué d'une interface couplée à une base de données dans laquelle sont stockées un ensemble de génomes procaryotes. Cette base de données contient aussi des informations complémentaires sur les gènes et les protéines associées, ainsi que sur les utilisateurs du site (voir **Annexe**).

L'objectif de notre outil est de permettre à tous les utilisateurs d'accéder aux informations relatives aux génomes contenues dans la base. De plus, notre outil permet à certains utilisateurs (voir partie II) d'annoter des génomes, dont les informations sont ensuite intégrées dans la base de données et rendues accessibles à l'ensemble des utilisateurs du site.

Nous présentons dans ce document les objectifs ainsi que le fonctionnement de l'outil du point de vue de l'utilisateur à partir d'une maquette de site web. Nous détaillons les fonctionnalités page par page.

I-Page d'accueil « ind.html » (Fig. 1, Fig. 2)

La page d'accueil de notre site est très simple : le visiteur peut choisir de se connecter ou s'inscrire en cliquant sur les boutons correspondants. S'il choisit de se connecter, un formulaire de connexion apparaît (adresse mail en identifiant, et mot de passe) et en validant, il accède alors à la page centrale (voir III). Si le visiteur n'a pas de compte, il peut en créer un en cliquant sur « *Don't have an account?* » ou « *Sign in* ». Il accède alors à la page de création de compte (voir II).

II-Page de création de compte « signup.html » (Fig. 3)

Le visiteur qui a sélectionné « *Don't have an account?* » ou cliqué sur « *Sign in* » sur la page d'accueil (voir I) accède à la page de création de compte.

Sur cette page, il peut créer un compte en remplissant un formulaire : prénom, nom, adresse email, numéro de téléphone, mot de passe, (avec confirmation) et choix du rôle.

L'utilisateur peut choisir parmi 3 rôles : lecteur, annotateur et valideur. Un lecteur peut lire les informations relatives aux génomes contenues dans la base. Un annotateur peut faire tout ce que fait un lecteur et en plus annoter des séquences. Un valideur peut faire tout ce que fait un annotateur et en plus valider des annotations des annotateurs.

Une fois le formulaire de création de compte rempli, le compte du visiteur est créé et ses informations sont ajoutées à la base de données. Il accède ensuite à nouveau à la page d'accueil (voir I), d'où il peut se connecter.

III-Page centrale « pageform.html » (Fig. 4)

Une fois connecté, le visiteur peut accéder à la page centrale du site. Cette page permet à tous les visiteurs d'accéder à toutes les fonctionnalités du site, en cliquant sur les deux premiers « boutons » qui apparaissent en haut à gauche.

-Le bouton « formulaire génome » redirige vers la page formulaire génome (voir IV)

-Le bouton « formulaire gène/protéine » redirige vers la page formulaire gène/protéine (voir V)

Les trois boutons qui suivent concernent le rôle du visiteur. Ces boutons sont tous présentés dans notre maquette à des fins explicatives, mais dans l'outil fonctionnel, ils ne seront présentés qu'aux visiteurs concernés :

- seuls les annotateurs, validateurs et administrateur verront le bouton « Annotator», qui redirige vers la [page annotateur](#) (voir VII)
- seuls les validateurs et administrateurs verront le bouton « Validator», qui redirige vers la [page validateur](#) (voir VI)
- seul l'administrateur verra le bouton « Administrator », qui redirige vers la [page administrateur](#) (voir VIII)

IV-Page formulaire génome « genomeform.html » (Fig. 5)

La page formulaire génome contient un formulaire à partir duquel tous les visiteurs du site peuvent effectuer des recherches pour visualiser des génomes selon plusieurs critères. Les champs du formulaire sont «GenomeName », «Serotype », «Species », «Nucleotide sequence », «Strain », «Chromosome Number », «Sequence size » et «Reference ». Ces critères sont combinés par un « et » : seuls les résultats respectant toutes les conditions sélectionnées seront affichés sur la [page résultat génome](#) (voir IX). L'utilisateur peut remplir autant de champs qu'il veut.

V-Page formulaire gène/proteine « GeneProtform.html » (Fig. 6)

La [page formulaire gène/proteine](#) contient un formulaire à partir duquel tous les visiteurs du site peuvent effectuer des recherches pour obtenir des informations sur des gènes et leurs protéines associées (notre outil ne concernant que des espèces procaryotes, nous nous limitons à des cas où une seule protéine correspond à chaque gène) selon plusieurs critères. L'utilisateur peut alors remplir de nombreux champs («AccessionNumber », «GeneSequence», «ProteineSequence», etc.). Ici aussi, ces critères sont combinés par un « et » : seuls les résultats respectant toutes les conditions sélectionnées seront affichés sur la [page résultat gène protéine](#) (voir X). L'utilisateur peut remplir autant de champs qu'il veut.

VI-Page validateur « Validatorpage.html » (Fig. 7)

La [page validateur](#) contient deux tables. Dans la première table « Genome for annotation », chaque ligne correspond à un génome à attribuer. Le nom du génome, l'espèce et la souche sont précisés. Dans la colonne « Annotator name », le validateur renseigne l'identité de l'annotateur à qui il attribue la séquence à annoter. Nous avons simplifié notre maquette en attribuant un génome à un annotateur : dans l'outil fonctionnel, nous pourrions attribuer des CDS plutôt que des génomes complets. Les annotateurs peuvent ensuite voir les séquences qui leur ont été attribuées depuis la [page annotateur](#) (voir VII).

Dans la seconde table de la [page validateur](#) « Validation of annotator suggestions », le validateur peut télécharger les propositions d'annotation des annotateurs, et décider ou non de les valider. Si elles sont validées, elles seront visibles comme séquences validées par l'ensemble des utilisateurs.

VII-Page annotateur « Annotatorpage.html » (Fig. 8)

La [page annotateur](#) contient une table dans laquelle chaque ligne correspond à un génome à annoter attribué par un validateur. Dans l'outil fonctionnel, les annotateurs auront donc chacun une table différente en fonction des génomes qu'ils auront à annoter. Ici aussi, nous avons simplifié notre maquette en attribuant un génome à un annotateur : dans l'outil fonctionnel, nous pourrions attribuer des CDS plutôt que des génomes complets. Pour chaque génome à annoter, le nom du génome, l'espèce et la souche sont précisés. L'annotateur a par ailleurs la possibilité de télécharger les fichiers à annoter. Il peut ensuite

soumettre ses fichiers annotés dans la colonne « annotation ». Les suggestions d'annotation ainsi proposées sont soumises aux validateurs dans la **page validateur** (voir VI).

VIII-Page administrateur « adminpage.html » (Fig. 9)

La **page administrateur** contient une table « User list » dans laquelle chaque ligne correspond à un utilisateur. L'administrateur peut voir les informations des différents utilisateurs, ainsi que la date et l'heure de leur dernière correction. Les utilisateurs qui se sont connectés récemment apparaissent en haut de la table. Pour chaque utilisateur, l'administrateur peut supprimer ou modifier le rôle de l'utilisateur.

IX-Page résultat génome « genome.html » (Fig. 10)

Une fois que l'utilisateur a bien rempli le formulaire génome (décrit en IV) et a cliqué sur le bouton soumettre il y a deux possibilités :

Dans le cas où il n'y a qu'un seul résultat possible: il sera redirigé vers une nouvelle page qui contient toutes les informations sur le génome qu'il a choisi (le nom du génome, le sérotype, l'espèce, la séquence nucléotidique, la souche, le nombre de chromosome et des plasmides, la taille de séquence nucléotidique du génome et ainsi si ce génome est annoté ou non). Dans le cas où il est annoté, le mail de l'annotateur sera affiché et un commentaire sur cette annotation écrit par le validateur sera également affiché. Dans le cas où le génome n'a pas été annoté, il y a juste le champ « Annotator statuts » qui sera affiché en mentionnant "Not annotated".

Dans le cas où plusieurs résultats sont possibles : L'utilisateur va être redirigé vers une page contenant la liste des résultats possibles, cette liste est cliquable, une fois qu'il clique sur un élément de cette liste, il va être redirigé vers une nouvelle page contenant des informations sur le résultat du génome choisi par l'utilisateur (la même page obtenue dans le cas d'un seul résultat). Ici, on a représenté la page dans le cas où il n'y a qu'un seul résultat. Dans l'outil final, le génome sera cliquable. Cliquer sur un gène redirigera vers la page résultat gène/proteine correspondante.

X-Page résultat gène/proteine « GeneProt.html » (Fig. 11)

Enfin, l'utilisateur accède à la page des résultats gène/proteine, après avoir effectué une requête en remplissant les informations dont il dispose dans le formulaire (voir V). Il peut ainsi accéder aux informations sur les gènes et les protéines contenues dans la base de données et notamment lire les séquences des gènes/protéines. L'utilisateur peut également se rendre sur plusieurs banques de données (génomiques, protéiques, procaryote et d'alignement génomique). Il peut télécharger les résultats de cette page en cliquant sur le bouton "download results" sous forme de fichiers ".txt".

Annexe : Base de données

Nous avons débuté le projet par la modélisation de l'application en vue de comprendre les éléments fondamentaux qui vont être pris en compte lors de la conception de l'application. Nous avons ainsi créé un diagramme de classes UML décrivant les différentes tables et associations de notre base. Le but étant de créer une base de données optimisée et cohérente. Nous avons opté sur une base de données avec 3 tables : Utilisateur, Génome et Gène/Protéine.

La table utilisateur permettra de stocker toutes les données relatives aux utilisateurs du site on y retrouve deux clés primaires : l'email et le mot de passe requis au moment de l'authentification, ainsi que leurs informations personnelles (nom, prénom, numéro de téléphone), la date et l'heure de leur dernière connexion et le rôle qu'ils endossent au sein du site. Cette table est reliée à la table Génome via une relation "Annot" (pour Annotation). Un annotateur peut annoter plusieurs génomes, mais un génome ne peut être annoté que par un seul utilisateur. On retrouve ainsi une clé étrangère "EmailAnnot" renvoyant vers la

table utilisateur. Les clés primaires de la table Génome sont le nom du génome, l'état d'annotation (non annoté, non validé, annoté et validé). Enfin, la dernière table Gène_Prot est reliée à la table génome via l'association "contient" (et donc également relié à la table utilisateur), afin de stocker pour chacun des génomes l'ensemble des séquences codantes le composant et leurs protéines associées.

Nous avons rencontré des difficultés lors de l'insertion des clés étrangères, en effet certaines clés primaires n'étaient pas considérées comme unique par PostgreSQL. Ce qui explique pourquoi, nous avons choisi de forcer l'unicité des clés primaires EmailAnnot et NomGenome avec la commande Unique.