

Projet : création d'un site WEB

L'objet de ce projet est de faire travailler des étudiants sur la création d'un site WEB en mode collaboratif.

Cahier des charges :

Ce site devra porter sur l'un de ces deux thèmes:

Cancérologie - Etude des fusions de gènes

Cinéma de Science-fiction (jusqu'à 1970) faisant intervenir l'Intelligence Artificielle

Ce travail fait partie du contrôle continu (40% de la note pratique du C2I). Cette note est conservée à la deuxième session. Ce projet s'effectue par groupe constitué de 3 (très exceptionnellement 2) binômes.

Chaque groupe choisit un thème unique (Cancérologie ou cinéma) et chaque binôme traitera une maladie (ou un film) différente.

Les pages HTML pourront être créées à l'aide d'un éditeur de votre choix mais nous avons une préférence pour les projets écrits directement en HTML (un bonus de points est accordé dans ce cas). L'usage de feuilles de style est autorisé si tous les membres du groupe de projet en maîtrisent le fonctionnement.

Dans tous les cas votre site devra contenir :

Une première page (page de présentation) mentionnant les personnes composant les 3 binômes. Un lien hyper texte est associé à chaque binôme. En cliquant sur ces liens, on aboutit sur les pages spécifiques des binômes.

Chaque binôme doit développer plusieurs pages :

- une page contenant des liens hypertexte vers les fichiers archives (zip, rar...) téléchargeables des TP complets (Word, Excel) effectués en salle et complément à faire au domicile (fichiers de chaque membre du binôme).

- Sur 4 pages maximum, le contenu relatif au thème unique choisi (Etude des fusions de gènes ou film, 1 par binôme)

➡ Pour l'ensemble des deux thèmes

1) Une mise en forme de qualité qui doit être homogène par groupe (police, choix des couleurs, mise en page, ...). Il s'agit d'une des principales difficultés de votre projet.

2) Nous testerons votre projet sur une version récente de FireFox ou Chrome, sur 2 résolutions d'écran différentes (1024x768 et 1680x1050).

3) Si vos pages sont trop longues pour être affichées entièrement à l'écran, il faudra y placer des points d'ancrages pour revenir en haut de la page ou atteindre facilement les différentes parties.

Hébergement :

Votre site devra être hébergé gratuitement (ce qui a toujours été le cas pour les années précédentes), vous serez alors noté sur 20, (Cf cours Html, pas d'aide sur ce point, c'est une des difficultés de votre projet, pour autant, d'importants éléments de réponses seront transmis en cours).

Vous nous transmettez les informations relatives à votre projet (lien hypertexte ou structure et nom des personnes de votre groupe) sur Moodle via la section "Informations relatives au site réalisé".

Dans le cas où vous n'avez pas réussi à faire héberger gratuitement votre site (cela n'est jamais arrivé les années précédentes), vous devrez nous transmettre la structure de votre site (fichiers html zippés, la première page de votre site doit s'appeler impérativement index.html) et votre note de projet ne pourra dépasser 12/20.

Aide : Deux séances de questions réponses seront éventuellement prévues avec inscription préalable (les modalités vous seront communiquées ultérieurement).

Choix du projet

Le choix de l'item pour chaque binôme s'effectue sur la plateforme Moodle dans votre ENT :

Dans la rubrique C2i, la section "Choix du sujet" vous amène à la liste des items sélectionnables.

Cette section est ouverte du Mercredi 8 février au dimanche 25 février.

- Les membres de chaque binôme doivent voter (pour le même item). Ce choix n'est plus modifiable après la fin de période de choix.
- Les 3 binômes d'un projet doivent choisir 3 maladies ou 3 films différents; un projet est exclusivement soit maladie soit film.
- Chaque item ne peut être choisi que par 9 binômes (de groupes différents) au maximum, sauf le cinéma (18 binômes).
- Le "verrouillage" de la disponibilité des items s'effectue automatiquement au fur et à mesure des choix : "Premier Arrivé Premier Servi"

Constitution du groupe de projet

Le choix du groupe de projet pour chaque binôme s'effectue sur la plateforme Moodle dans votre ENT : Dans la rubrique C2i, la section "Constitution des groupes de projets" vous amène à la liste des items sélectionnables.

Cette section est ouverte du Mercredi 8 février au dimanche 25 février.

Sélectionnez le groupe/projet dans lequel vous voulez intégrer votre binôme.
Chaque groupe/projet est constitué par 3 binômes, (ou 2 binômes + 1 monôme) donc 6 (ou 5) personnes.
Attention, les 6 membres (ou 5 membres s'il y a un monôme) doivent voter pour le même groupe de projet

La date de fin de projet est fixée au 15/04/2017 minuit.

Pour toute demande d'information, veuillez utiliser impérativement votre courriel universitaire xxx@etu.umontpellier.fr. Les courriels émanant de mails personnels sont automatiquement rejetés.

Bon courage

Patrice Ravel, Guilhem Kister, Jean Peyhardi, Ali Jembain, Emmanuel Cornillot

Voici les spécifications par thème :

Projet Web C2I 2017-2018
Cinéma de Science-fiction (jusqu'à 1970) faisant intervenir l'Intelligence Artificielle
Guilhem Kister (guilhem.kister@umontpellier.fr)

Contenu minimal :

- ✓ Titre original et titre français
- ✓ Année de production, année de distribution (originale, et en France)
- ✓ Affiche originale et affiche française
- ✓ Réalisateur (avec photographie et filmographie)
- ✓ Acteurs (avec photographie et filmographie pour les principaux)
- ✓ Scénariste (avec photographie et filmographie)
- ✓ Compositeur (avec photographie et filmographie)
- ✓ Caractéristiques techniques : durée, format écran, format pellicule, technologie audio...
- ✓ Synopsis et résumé du film
- ✓ Localisation du tournage, Photos du tournage
- ✓ S'il s'agit d'une adaptation d'une œuvre littéraire :
 - Titre de l'œuvre (original, français), année
 - Auteur + bibliographie sommaire
- ✓ S'il s'agit d'un film élément d'un ensemble, informations sur ce dernier et les autres films le constituant (titres, dates, réalisateur, synopsis, nom des principaux acteurs)
- ✓ Lien vers page web d'un site cinématographique traitant du film.
- ✓ Extrait du film soit via un lien externe, soit directement inclus dans votre page.

Projet Web C2I 2017-2018
Etude des fusions de gènes
Pr. Emmanuel Cornillot (emmanuel.cornillot@umontpellier.fr)

Le but de l'étude est de caractériser des fusions gènes répertoriées dans la base de données « Tumor fusion gene data portal » du Jackson Laboratory <http://www.tumorfusions.org/>

Les gènes codent pour des protéines qui portent différentes fonctions biologiques et qui sont adressées à différent compartiment de la cellule. Dans le cas d'une fusion de gènes, le produit des deux gènes est maintenant fusionné au sein d'une même protéine.

Les informations sur les gènes de la fusion et les fonctions associées aux protéines codées par ces gènes sont référencées dans des bases de données accessibles publiquement sur internet.

L'objectif est de produire un site web qui fournisse les informations disponibles sur des protéines de fusions trouvées dans un cancer.

Le cancer est une pathologie liée au génome.

Une fusion de gènes apparaît suite à une délétion ou une inversion au sein d'un chromosome ou une translocation entre deux chromosomes.

Vous devez choisir deux exemples de fusion de gènes au sein d'un même cancer incluant dans un premier cas un gène codant pour une protéine avec une activité enzymatique de type kinase ou phosphatase et dans l'autre cas une protéine avec un domaine membranaire.

Attention il est possible de trouver des combinaisons de gènes respectant les deux conditions, mais vous devez, dans la mesure du possible, traiter le cas d'une protéine de fusion soluble et une protéine membranaire (voir ci-dessous).

Etape 1 : Choix du cancer que vous souhaitez analyser

La base de données TCGA met à disposition les données génétiques de plus de 10000 patients atteints de différents cancers.

1.1. Choix du cancer

Les abréviations décrivant les cancers sont données à la page suivante

<https://gdc.cancer.gov/resources-tcga-users/tcga-code-tables/tcga-study-abbreviations>

Aller sur le site du Jackson laboratory <http://www.tumorfusions.org/>

1.2. Recherche les fusions par type de cancer (Cancer type)

Sélectionner en haut le cancer (commencer les analyses avec tier1).

Le site vous fournit une série de fusion de gènes

TCGA SampleID correspond à l'identifiant du patient dans lequel la fusion a été détectée

Vous devez choisir les fusions parmi celle qui sont marquée « In-frame ».

Etape 2 : Sélection des fusions à analyser

Les fusions de gènes sont à chercher dans les jeux de données tier1 ou tier2 seulement.

Avec la souris, sélectionner une fusion de gènes In-frame

Les deux partenaires de la fusion sont séparés par un double tiret bas (ex : FGFR2__BICC1). Il s'agit de l'identifiant du gène appelé Hugo symbol que vous devez copier pour chercher les informations dans les bases de données Ensembl (<http://www.ensembl.org>) ou Gene Cards (<http://www.genecards.org>).

La base de données du Jackson Laboratory vous donne aussi les informations sur l'orientation des gènes : /1, brin direct ou /-1, brin complémentaire. Si les gènes sont sur deux chromosomes différents, il s'agit d'une translocation. S'ils sont sur un même chromosome et dans le même sens /1 ; /1 la fusion est le résultat d'une délétion, si les gènes sont en sens inverse /1 ; /-1 ou /-1 ; /1, il s'agit d'une inversion. Les trois événements sont à prendre en compte dans votre étude.

Vous devez chercher deux types de fusion. La première fusion impliquera deux protéines solubles. Une protéine soluble est une protéine qui n'est pas ancrée à la membrane par un domaine transmembranaire ou une ancre GPI. La deuxième fusion devra impliquer au moins une protéine membranaire (GPI ou transmembranaire).

La présence d'un ou plusieurs domaines transmembranaires est aussi fournie par la base de données Pfam (<http://pfam.xfam.org>). Vous devez regarder la description des domaines protéiques dans la base de données Pfam et chercher la présence de « transmembrane ». Il faut utiliser l'identifiant UNIPROT de votre protéine pour chercher dans Pfam (barre latérale gauche, Jum to – enterID/acc). L'identifiant UNIPROT est donné par la base de données Gene Cards (ex : FGFR2 -> [P21802](#))

Il faut qu'au moins une des deux protéines de la première fusion impliquant deux protéines solubles ait une activité kinase ou phosphatase.

Il n'y a pas de prédiction pour un ancrage GPI dans la base de données Pfam. Vous devez utiliser les informations fournies sur la protéine, sa fonction et sa localisation par les bases de données Gene Cards et The Human Protein Atlas (<https://www.proteinatlas.org>).

Etape 3 : Eléments d'information à fournir dans le site web

Les informations fournies dans le site web doivent être sous forme de lien hypertexte. Vous analyserez les deux fusions l'une après l'autre. Si vous ne trouvez pas d'exemple pour l'une ou l'autre des conditions, alors vous choisirez une fusion arbitrairement. N'oubliez de donner une conclusion générale sur le rôle potentiel de la protéine de fusion obtenue.

Information sur le cancer

Cette partie du site sert d'introduction et doit donner des informations générales sur le cancer étudié et renvoie aux sites principaux sur cette pathologie.

Composition de la cohorte TCGA pour le cancer choisi.

Nb de patients analysé

Gènes fréquemment mutés

Données génétiques générales sur ce cancer

COSMIC <http://cancer.sanger.ac.uk/cosmic>

cBioPortal <http://www.cbioportal.org/>. Choisissez le projet marqué (TCGA, Provisional) pour obtenir les informations sur la cohorte analysée par le Jackson Laboratory.

Les associations de malades ou sites d'information sur la maladie.

<http://www.cancer-genetics.org/clink30.htm>

Orphanet, Internet en général

Site de l'INCA pour le cancer (<http://www.e-cancer.fr/>)

Les gènes impliqués dans la fusion

Bases de données

ENSEMBL : http://www.ensembl.org/Homo_sapiens/Info/Index,

Gene Cards : <http://www.genecards.org/>

Informations à fournir

Avoir des informations sur les gènes

Sur quel chromosome

Combien intron/exon

Sens de transcription sur génome de référence

Description des cDNA

Nombre d'isoforme

Expression différentielle en fonction des tissus

Les protéines impliquées dans la fusion

Bases de données

Gene Cards et The Human Protein Atlas (<https://www.proteinatlas.org>) pour la description de la fonction.

UNIPROT (<http://www.uniprot.org>). L'identifiant UNIPROT est récupéré au niveau de la base de donnée Gene Cards. Il est possible de récupérer la séquence de la protéine au format FASTA.

Pfam (<http://pfam.xfam.org>). Utiliser l'identifiant UNIPROT pour trouver la page de la protéine. Une protéine est membranaire si elle possède un domaine appelé « transmembrane »

Analyse complémentaires (pour ceux qui souhaitent faire un peu de bioanalyse)

Ces analyses ne sont pas obligatoires.

Il est possible de prédire certains domaines protéiques en utilisant la séquence au format FASTA.

Peptide signal : SignalP, <http://www.cbs.dtu.dk/services/SignalP/>

Adressage cellulaire : TargetP, <http://www.cbs.dtu.dk/services/TargetP/>

Domaine transmembranaire : TMPred, https://www.ch.embnet.org/software/TMPRED_form.html

Ancrage GPI : PredGPI, <http://gpcr.biocomp.unibo.it/predgpi/>

Vous pouvez chercher d'autres sites d'analyse de séquences protéiques si vous le souhaitez.

Relation avec les maladies héréditaires

OMIM (<https://www.omim.org>). Effectuer la recherche par mots clés sur le cancer et/ou de la fusion de gènes (ex : ETV6 NTRK3 fusion).

Certains cancers peuvent avoir une origine héréditaire. Certaines fusions de gènes ont aussi été associées à des maladies héréditaires.

Conclusions sur le rôle de la fusion

Bibliographie scientifique

PubMed (<https://www.ncbi.nlm.nih.gov/pubmed/>)

Revue d'internet

Equipes de recherche travaillant sur le sujet

Autres informations