

Editorial



Bien qu'ils n'étudient pas tous la même maladie, le même virus ou le même insecte, tous les chercheurs ont besoin de recueillir et de stocker une quantité considérable de données et de comparer leurs résultats avec les connaissances de leurs pairs. En d'autres termes, ils doivent pouvoir accéder à des données et logiciels de tout premier ordre. Le SIB Institut Suisse de Bioinformatique soutient la communauté des sciences de la vie à bien des égards. Il développe un large éventail de ressources telles que SWISS-MODEL, créée il y a près de vingt ans et aujourd'hui essentielle pour les chercheurs du monde entier, UniProtKB/Swiss-Prot, STRING,

EADock, Melanie ou encore neXtProt, la plateforme dédiée aux protéines humaines qui sera mise en ligne dans quelques jours. Le SIB participe également à des projets de recherche européens en mettant son expertise au service de la lutte contre des maladies telles que le diabète. Le SIB a toujours su que la collaboration est primordiale et qu'elle doit être renforcée, tant sur le plan interne que sur le plan externe. Les groupes de recherche et de service du SIB collaborent ainsi très étroitement en vue de relier leurs différentes bases de données et plateformes de logiciels. Cette interopérabilité ouvre aujourd'hui une nouvelle dimension aux ressources mises à disposition des chercheurs en sciences de la vie.

Pour garantir l'avenir et le progrès de la recherche, la seule création de telles ressources n'est cependant pas suffisante. Il faut aussi pouvoir assurer leur fiabilité à partir de deux critères essentiels: la qualité et la durabilité. Quel soutien représenterait une base de données susceptible de disparaître d'un jour à l'autre? Des années de travail viendraient ainsi à disparaître en un clin d'œil. Il est de ce fait primordial de créer une infrastructure solide pour sécuriser l'avenir des ressources bioinformatiques à l'échelle internationale. C'est précisément ce but que s'est fixé l'Union Européenne dans le cadre de son projet ELIXIR. Le SIB est fier de contribuer à cet effort européen et impatient de voir les résultats et l'impact de cette initiative.

Nous espérons que vous aurez plaisir à lire ce nouveau numéro de *Swiss Bioinformatics*. Vous y découvrirez un panel de nos activités sur la scène nationale et internationale de la bioinformatique.

Ron Appel, Directeur

Les Brèves

Subvention NIH pour UniProt

Les NIH (National Institutes of Health, Etats-Unis) ont accordé une subvention de trois ans au consortium UniProt, qui inclut le SIB, l'EBI (European Bioinformatics Institute) et le PIR (Protein Information Resource). Le SIB bénéficiera ainsi d'un soutien financier de plus de deux millions de francs par an pour la maintenance et l'optimisation de la base de données UniProtKB/Swiss-Prot. Cette subvention souligne l'importance et le rayonnement international du SIB et de ses groupes de service et de recherche.

neXtProt, une nouvelle étape

La première version publique de neXtProt, la nouvelle plateforme de connaissances et de logiciels dédiée aux protéines humaines, sera mise en ligne dans quelques jours. neXtProt devrait devenir une ressource essentielle pour les chercheurs dans le domaine biomédical qui s'intéressent au rôle des protéines dans la santé et les maladies. neXtProt a été développée conjointement par le SIB et GeneBio. Cette collaboration est essentielle car elle garantit une prise en compte des besoins de tous les chercheurs concernés tant dans le domaine académique que dans l'industrie. neXtProt repose sur l'excellent travail d'annotation réalisé par le groupe Swiss-Prot et prévoit des synergies avec d'autres ressources développées par le SIB. Cette première version utilise déjà des données et analyses encapsulées dans la base de données Bgee, que le groupe «Evolutionary Bioinformatics» du SIB développe à l'Université de Lausanne.

Si cette première version n'est qu'un petit aperçu du large spectre de données et outils que neXtProt mettra à disposition des utilisateurs, elle n'en est pas moins une étape importante dans la réalisation de l'une des missions du SIB: la création de ressources de qualité pour les chercheurs en sciences de la vie.

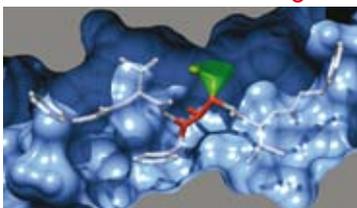
Le serveur neXtProt, hébergé par le groupe Vital-IT, est disponible sous: beta.nextprot.org.

● Novembre 2010

Table des matières

| | |
|--|---|
| ● Editorial | 1 |
| ● Les Brèves | 1 |
| ● Recherche | 2 |
| ● Collaboration internationale | 2 |
| ● Formation et sensibilisation | 3 |
| ● Lien avec l'industrie | 3 |
| ● A propos de la bioinformatique et du SIB | 4 |

EADock désormais en ligne



Les protéines dysfonctionnelles peuvent causer de nombreuses maladies, y compris des cancers. Le rôle de nombreux médicaments anti-cancéreux consiste à moduler l'activité de ces protéines. Ces médicaments sont de petites molécules conçues pour venir se loger dans le site actif de la protéine afin de stimuler ou d'inhiber une fonction bien précise, un peu comme une clé qui permettrait d'ouvrir ou de verrouiller une serrure.

La prévision de telles interactions à l'échelle atomique et dans un espace tridimensionnel a été l'un des principaux sujets de recherche du groupe «Molecular Modelling» du SIB au cours des dix dernières années. Le fruit de ces recherches prend aujourd'hui la forme d'un programme informatique appelé EADock. Il permet l'identification de nombreuses molécules agissant contre des protéines cibles impliquées dans divers types de cancers.

Le savoir-faire concentré dans ce logiciel est aujourd'hui à la disposition de toute la communauté scientifique. Le logiciel peut être désormais utilisé en ligne gratuitement via le site SwissDock^(*). Dans les années à venir le SIB entend étendre son utilisation au-delà du cercle des experts dans le domaine et permettre ainsi des progrès significatifs en oncologie moléculaire et plus généralement dans la recherche pharmaceutique.

(*) www.swissdock.ch

Unis contre le diabète

Pour lutter contre la pandémie du diabète, la biotechnologie et l'industrie pharmaceutique ont uni leurs efforts dans le cadre d'IMIDIA, un projet de l'Innovative Medicines Initiative (IMI) centré sur le diabète. Le projet dirigé par le Professeur Bernard Thorens de l'Université de Lausanne a été lancé officiellement le 14 juin 2010. Il devra permettre le développement de biomarqueurs et d'outils qui permettront une



meilleure prise en charge du diabète et, à plus long terme, un traitement de cette pathologie.

Avec plusieurs de ses groupes de recherche et de service impliqués dans le projet, le SIB apporte plusieurs niveaux de support: bioinformatique, informatique et analyse des données.

Recherche

SWISS-MODEL et STRING: des services à rayonnement international

Le corps humain contient environ cinq millions de protéines produites par 20'000 gènes différents. Nul doute que l'identification de l'ensemble de leur rôles et interactions continuera à occuper la communauté des chercheurs pendant de nombreuses années. L'une des missions du SIB consiste à soutenir les chercheurs en mettant à leur disposition des bases de données et des logiciels centrés sur les protéines. Certains services développés par notre institut sont devenus des outils essentiels pour les scientifiques du monde entier.

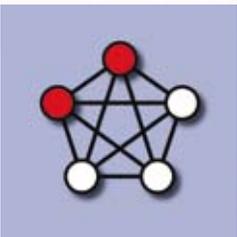


SWISS-MODEL est un service qui permet de modéliser des structures de protéines en trois dimensions. Les biochimistes et biologistes moléculaires du monde entier utilisent cet outil et en apprécient l'accès gratuit et la facilité d'emploi. Les utilisateurs n'ont qu'à entrer une séquence d'acides aminés ou le code d'accès UniProt d'une protéine pour construire son modèle 3D. Outre cette méthode automatisée

de modélisation, SWISS-MODEL offre plusieurs options qui permettent aux utilisateurs spécialisés de préciser certains paramètres, tels que la sélection du modèle structurel le plus approprié. SWISS-MODEL Repository fournit des modèles pré-calculés de protéines humaines ainsi qu'une sélection d'organismes modèles, ce qui facilite l'échange de données avec d'autres services comme Swiss-Prot et STRING. La modélisation des structures protéiques trouve de nombreuses applications dans le domaine de la recherche biomédicale, telles que la conception de médicaments basée sur la structure, la planification rationnelle d'études fonctionnelles, ou l'ingénierie des protéines.

Depuis son lancement en 1993, SWISS-MODEL n'a cessé d'être amélioré. En 2001, son développement a été repris par le groupe «Structural Bioinformatics» de Torsten Schwede du Biozentrum (Université de Bâle) et du SIB. Chaque année, plus de 200'000 utilisateurs dans le monde accèdent au serveur et visualisent un demi-million de protéines en 3D.

SWISS-MODEL est accessible à l'adresse suivante: swissmodel.expasy.org.



STRING

L'identification et la modélisation de structures ne sont qu'une première étape vers la compréhension de toutes les fonctions d'une protéine. La situation est bien plus complexe, les protéines ayant une «vie sociale». En effet, chaque protéine interagit de diverses manières avec ses congénères. Ces interactions peuvent se traduire par des changements considérables

dans le comportement d'une protéine, comme par exemple l'acquisition d'une nouvelle fonction. Actuellement on dispose de nombreuses données sur les interactions protéine-protéine, souvent dispersées dans plusieurs ressources publiques. Pour les chercheurs, il serait souhaitable de disposer d'un système unique qui intègre et ré-évalue les informations et en facilite la recherche et la consultation. C'est précisément ce service que propose la base de données STRING (Search Tool for the Retrieval of Interacting Genes), qui regroupe toutes les informations sur les liens fonctionnels entre les protéines. Les atouts prin-

cipaux de STRING sont son exhaustivité, son indice de confiance et son interface très interactive et intuitive. STRING est le seul site à inclure plusieurs centaines d'organismes (et bientôt plus de 1'100) allant de la bactérie aux êtres humains.

Créée en 2000 à l'EMBL à Heidelberg (Laboratoire Européen de Biologie Moléculaire) la base de données STRING est développée et maintenue aujourd'hui par un petit consortium de groupes de bioinformatique en Europe. Le SIB joue un rôle primordial dans le développement de cette ressource à laquelle participe un nombre important de collaborateurs du groupe de Christian von Mering (Université de Zurich et SIB). STRING est utilisée dans le monde entier et de nombreux scientifiques la téléchargent et l'utilisent hors ligne, ce qui en fait une ressource clé dans les projets de recherche internationaux.

STRING est accessible à l'adresse suivante: string-db.org.

UNE INTEROPERABILITE ACCRUE

Avec 29 groupes de recherche et de service basés dans les différentes institutions et villes suisses, le SIB est un modèle en termes de collaborations inter-institutionnelles et inter-cantoniales. Aujourd'hui, le SIB renforce cet aspect collaboratif en reliant les différentes ressources et services qu'il développe et maintient pour la communauté internationale des chercheurs en sciences de la vie. Concrètement, cela signifie que le SIB veut donner aux utilisateurs la possibilité de travailler avec plusieurs ressources simultanément pour une plus grande convivialité et efficacité. STRING et SWISS-MODEL ont déjà fait un grand pas dans cette direction. Leurs utilisateurs peuvent en effet visualiser une structure de protéine et ouvrir à tout moment une fenêtre avec des informations issues de la base de données STRING et vice-versa. Nos services atteignent ainsi une nouvelle dimension et gagnent en efficacité pour un plus grand soutien à la communauté des chercheurs.

Collaboration internationale

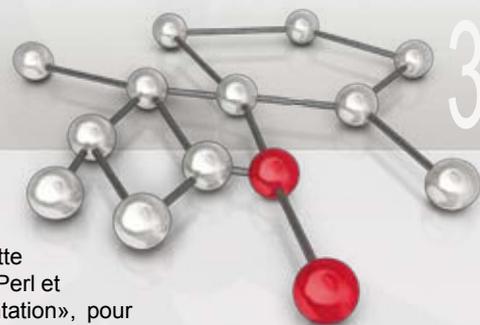
ELIXIR de vie

Comment les chercheurs en sciences de la vie pourraient-ils poursuivre leurs travaux sans l'accès à toutes les connaissances rassemblées à ce jour dans les bases de données? Comment pourraient-ils, par exemple, identifier un virus si les bases de données contenant toutes les informations pertinentes, telles que les séquences des gènes, venaient soudainement à disparaître? L'espoir de pouvoir un jour prévenir ou guérir le cancer s'évanouirait au même moment.

La réponse à ce scénario alarmiste est ELIXIR - European Life Science Infrastructure for Biological Information, une initiative de l'Union Européenne visant à garantir l'avenir des données biologiques. Le premier objectif est de sécuriser le soutien financier des agences gouvernementales, organismes de bienfaisance, industries et organisations intergouvernementales en Europe. Le second objectif consiste à renforcer et pérenniser une infrastructure de classe mondiale pour la gestion et l'intégration des informations liées aux sciences de la vie. La réalisation de ces objectifs repose sur deux facteurs majeurs: la coordination et la garantie d'un financement durable.

A quoi mènera ELIXIR?

Le projet réunit tous les acteurs européens de la scène bioinformatique, des fournisseurs de données aux utilisateurs, afin d'évaluer les besoins et ressources disponibles. Il doit aboutir à la création d'un réseau bioinformatique structuré qui fournira à l'ensemble de la communauté des sciences de la vie les données, les outils, la coordination, l'intégration, le support, les standards et la formation dont elle a besoin. Ceci permettra ultimement une amélioration de la qualité de vie grâce à de meilleurs soins aux populations vieillissantes, à un approvisionnement alimentaire durable, à des industries biotechniques et pharmaceutiques compétitives et à une meilleure protection de l'environnement.



Comment la Suisse et le SIB contribuent-ils au projet?

En tant qu'infrastructure suisse des sciences de la vie pour les données biologiques, le SIB est déjà au niveau national ce qu'ELIXIR projette de devenir au niveau européen. Grâce au soutien de la Confédération, la Suisse est le seul pays doté d'une solide infrastructure bioinformatique. Et il ne fait aucun doute que le SIB, qui met des ressources de renom à la disposition de la communauté internationale, jouera un rôle important dans le projet ELIXIR.

Pour plus d'informations: www.elixir-europe.org.

Liens avec l'Inde renforcés



En 2009, le Conseiller en Sciences et Technologie de l'ambassade Suisse à New Delhi et le SIB ont décidé de réunir les bioinformaticiens des instituts suisses et indiens afin d'envisager de futures collaborations. Un symposium a ainsi été organisé à l'Institut Indien de Technologie IIT (Indian Institute of Technology) à Delhi durant la dernière semaine du mois d'octobre 2010. Le SIB et les principaux instituts indiens de bioinformatique y ont présenté leurs activités et projets. Les échanges ouverts qui ont eu lieu avec des représentants du gouvernement indien ont posé les jalons d'intéressantes collaborations futures.

Formation et sensibilisation

Le SIB dispense des cours de bioinformatique, organise des ateliers et joue un rôle de consultant auprès des chercheurs en sciences de la vie. Une forte demande pour ces cours est à signaler, tel qu'en attestent le taux élevé de participation et les listes d'attente. Plusieurs cours et ateliers ont été organisés cette année:

Quatre cours d'une semaine

- Introduction aux statistiques pour les biologistes
- Perl et techniques de présentation
- Analyse bioinformatique des données de séquençage UHT (Ultra-High-Throughput)
- Statistiques avancées: modélisation statistique

Cours d'un ou deux jour(s) et ateliers

- Programmation Perl pour les sciences de la vie
- Informatique Haute Performance (High-performance computing) en sciences de la vie
- Identification des protéines par spectrométrie de masse.

Toutes les sessions étaient complètes et les formations d'une journée ou deux ont même été doublées. La qualité des cours et la prise en compte des besoins actuels expliquent ce succès.



Le point culminant de cette année a été le cours sur «Perl et les techniques de présentation», pour lequel le SIB a invité Damian Conway, l'un des plus grands spécialistes du langage Perl dans le monde qui est venu spécialement d'Australie. Perl (Practical Extraction and Report Language) est un langage de programmation très populaire auprès des personnes ayant à écrire régulièrement des programmes de traitement de données. C'est par exemple le cas de Monica Albarca Aguilera, chercheuse post-doctorale en biologie à l'EPFL (Ecole Polytechnique Fédérale de Lausanne): «J'ai trouvé le cours sur Perl très intéressant. Il nous a donné un bon aperçu des possibilités qu'offre ce langage de programmation. Damian Conway nous a fourni des explications très claires et il a su captiver notre attention tout au long de la session.» Ayant besoin d'analyser les données à haut débit qu'elle génère dans son travail, Monica a déjà suivi plusieurs cours organisés par le SIB et espère qu'il y en aura d'autres.

Forts de ce succès, les membres de l'équipe formation du SIB ont déjà prévu le programme de 2011, qui proposera entre autres:

- un cours d'introduction aux statistiques
- un cours Unix/Perl
- un cours d'introduction à l'analyse de données microarrays
- un cours avancé de statistiques
- la reconduction de certains des ateliers organisés en 2010 mais aussi de nouveaux thèmes axés sur divers projets bioinformatiques.

Tous les cours de bioinformatique en un clic!

Promouvoir la bioinformatique et coordonner la formation en Suisse font partie des missions du SIB. Organisateur d'une grande partie des cours de bioinformatique, le SIB propose dès à présent une liste complète de tous les cours dispensés en Suisse. Cette liste est en ligne sur le site internet du SIB (www.isb-sib.ch/education.html). Avec ce nouveau service, le SIB veut faciliter l'accès des étudiants à l'information et leur inscription aux cours proposés. Les étudiants ont la possibilité de souscrire à une mailing-list qui les informe automatiquement lorsqu'un nouveau cours est proposé. Environ 1'500 personnes, principalement des participants de cours précédents, sont déjà inscrites et ce nombre pourrait continuer à augmenter à chaque nouvelle session.

Lien avec l'industrie

MELANIE, un logiciel qui fait la différence

MELANIE est un logiciel pour l'analyse d'images de gels d'électrophorèse bidimensionnelle. Sur ces gels, des protéines – contenues par exemple dans un échantillon de sang ou de tissu – sont séparées en fonction de leur taille et de leur charge. Le résultat est une sorte de cartographie en protéines de l'échantillon d'origine.

Le logiciel permet de comparer les gels provenant de divers échantillons et de mesurer les changements dans l'expression des protéines à des fins de statistiques et d'analyses. Ce type d'expérience est nécessaire lorsque l'on cherche des protéines permettant de distinguer les cellules saines des cellules cancéreuses ou que l'on veut évaluer l'efficacité d'un médicament sur un organisme.



L'objectif est d'identifier des protéines spécifiques qui permettent de distinguer des échantillons, ou d'étudier les motifs de co-expression de protéines afin de permettre

une meilleure compréhension de la régulation et des interactions protéiniques. Ces informations pourront alors être utilisées pour le développement de biomarqueurs pour les diagnostics thérapeutiques ou pour l'identification de médicaments.

Démarré en 1983 dans le cadre d'un projet de doctorat, Melanie est devenu un logiciel indispensable pour des milliers d'utilisateurs dans le monde entier depuis sa première version commerciale en 1994. Melanie est aujourd'hui développé et constamment amélioré par le groupe «Proteome Informatics» dirigé par Frédérique Lisacek, sous

l'œil avisé de Daniel Walther qui supervise depuis 1996 l'équipe de développement. L'équipe travaille en étroite collaboration avec Sonja Voordijk, responsable produit chez GeneBio, et avec GE Healthcare qui commercialise le logiciel sous le nom ImageMaster™ 2D Platinum. Les droits de licence provenant de cette coopération avec ces partenaires commerciaux permettent au SIB de financer les nouveaux développements du logiciel.

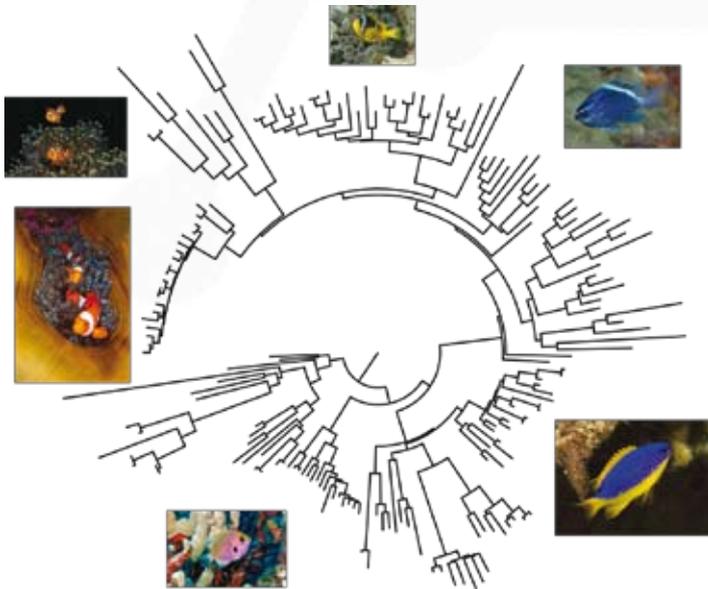
Les chercheurs des milieux pharmaceutiques, biotechnologiques et académiques ont recours à Melanie pour répondre à des questions telles que: «Peut-on identifier des protéines ou motifs de protéines caractéristiques d'une tumeur?». Les divers outils d'analyse disponibles dans Melanie guident et aident le chercheur à sélectionner les protéines qui seront soumises à des analyses complémentaires (notamment par spectrométrie de masse).

(ImageMaster est une marque déposée de GE Healthcare companies.)

Qu'est-ce que la bioinformatique?

Ces 30 dernières années, l'émergence de nouvelles techniques en biologie et les avancées en informatique ont augmenté à la fois la quantité et la complexité des données biologiques. C'est pourquoi les scientifiques doivent maintenant souvent appliquer les technologies de l'information pour résoudre des problèmes biologiques. Cette science est appelée *bioinformatique*.

Les biologistes se servent de la bioinformatique pour stocker, traiter et analyser de grandes quantités de données – ceci pour mieux connaître et mieux comprendre les processus biologiques. Ces connaissances peuvent à leur tour mener à des découvertes scientifiques permettant d'améliorer la qualité de vie – par exemple concevoir des traitements médicaux plus efficaces ou améliorer les rendements agricoles.



A propos du SIB www.isb-sib.ch

Le SIB Institut Suisse de Bioinformatique est une fondation académique à but non lucratif qui fédère les activités de bioinformatique en Suisse. Sa double mission est d'une part, d'offrir des ressources bioinformatiques de très haut niveau à la communauté des chercheurs dans le domaine des sciences de la vie au niveau national et international, dans des disciplines telles que la génomique, la protéomique et la biologie des systèmes, et d'autre part, de diriger et coordonner le domaine de la bioinformatique en Suisse.

L'institut est reconnu dans le monde entier pour son activité de développement de logiciels de pointe et de bases de données soigneusement annotées, dédiés à la recherche dans les sciences de la vie. Le SIB est composé de 29 groupes de recherche et de services de tout premier ordre, réunissant plus de 400 scientifiques bioinformaticiens dans les domaines de la protéomique, la transcriptomique, la génomique, la biologie des systèmes, la bioinformatique des structures, la bioinformatique de l'évolution, la modélisation, l'imagerie, la biophysique et la génétique des populations, répartis entre Bâle, Berne, Genève, Lausanne et Zurich. Le savoir-faire du SIB est reconnu au niveau international et ses infrastructures et ressources en bioinformatique sont utilisées par des chercheurs actifs dans le domaine des sciences de la vie dans le monde entier.

Membres institutionnels:

- École polytechnique fédérale de Lausanne (EPFL)
- École polytechnique fédérale de Zurich (ETHZ)
- Université de Bâle
- Université de Berne
- Université de Genève
- Université de Lausanne
- Université de Zurich
- Ludwig Institute for Cancer Research (LICR)
- Friedrich Miescher Institute for Biomedical Research (FMI)
- Geneva Bioinformatics (GeneBio) S.A.
- Hewlett Packard

Une liste complète de projets SIB est disponible sur: www.isb-sib.ch/research/projects.

Arbre phylogénétique
© Nicolas Salamin et Glenn Litsios



Swiss Institute of Bioinformatics

SIB | Swiss Institute of Bioinformatics

Quartier Sorge
Bâtiment Géopode
CH-1015 Lausanne
Suisse
t +41 21 692 40 50
f +41 21 692 40 55
www.isb-sib.ch

Swiss Bioinformatics est rédigé par la Communication du SIB avec la contribution des membres du SIB. Conception et mise en page: D. Meyer