

COLLOQUE BIO-INFORMATIQUE ROBERT-CEDERGREN BIOINFORMATICS COLLOQUIUM

Programme

Université de Montréal
2-3-4 novembre 2005

Présentations orales et d'affiches
Poster and oral presentations

www.bioinfo.umontreal.ca/evenements/colloquecedergren.html

Bienvenue au 2e colloque bio-informatique Robert-Cedergren !

Ce colloque se veut le rendez-vous annuel de la communauté universitaire oeuvrant en bio-informatique au Québec. L'objectif principal est de partager les derniers développements en ce domaine par le biais d'un concours d'affiches et de présentations orales et de rendre compte de l'importance grandissante de la bio-informatique dans les sciences de la vie.

En cette deuxième édition du colloque, les conférenciers invités seront Bernard Moret, professeur à l'Université du Nouveau-Mexique et directeur du CIPRES (Cyber Infrastructure for Phylogenetic Research) et Shoshana J. Wodak, professeur à l'Université de Toronto et titulaire de la Chaire de recherche du Canada en biologie computationnelle et en bio-informatique.

Dix-sept présentations orales et seize affiches seront en lice dans cinq catégories. L'Université ayant le meilleur classement remportera le prix Robert-Cedergren.

Cette sculpture d'acrylique représente l'outil de travail par excellence du bio-informaticien : l'ordinateur. Sur l'écran, une molécule d'ARN de transfert, thème de recherche cher à Robert Cedergren.



Les prix individuels seront décernés dans les catégories suivantes :

	Meilleures présentations orales	Meilleures affiches
2 ^e cycle	2000 \$	1000 \$
3 ^e cycle	2000 \$	1000 \$
Post-doctorant	--	--

Un excellent colloque bio-informatique à tous et à toutes !

B. Franz Lang, Ph.D.
Directeur
Centre Robert-Cedergren

Gertraud Burger, Ph.D.
Responsable des programmes
de 2e et 3e cycle – Bio-informatique

Welcome to the 2nd annual Robert Cedergren Bioinformatics Colloquium!

This second Colloquium is an annual event gathering the university community working in Bioinformatics in Quebec. The main purpose of this event is to share the latest Bioinformatics developments, by posters and oral presentations to take into account the increasing role of Bioinformatics in life sciences.

Keynote speakers will be Dr. Bernard Moret, professor at New Mexico University and Director of CIPRES (Cyber Infrastructure for Phylogenetic Research) and Dr. Shoshana Wodak professor at University of Toronto and holder of Canada Research Chair in Computational Biology and Bioinformatics.

This year, 17 oral presentations and 16 posters will compete in 5 categories. The university with the best ranking will win the Robert Cedergren Award.

This sculptur made of acrylic represents the bioinformatician's tool : the computer. On the screen is displayed, a tRNA molecule, one of the preferred research themes of Robert Cedergren.



Individual awards will be given in the following categories:

	Best oral presentations	Best posters
Ms	2000 \$	1000 \$
PhD	2000 \$	1000 \$
Postdoctoral students	--	--

Enjoy this 2nd annual Robert-Cedergren Bioinformatics Colloquium!

B. Franz Lang, Ph.D.
Director
Robert Cedergren Centre

Gertraud Burger, Ph.D.
Leader
Bioinformatics graduate programs

Comités/Committees

Comité d'organisation / Organizing Committee

Gertraud Burger
B. Franz Lang
Hervé Philippe
Marie Robichaud

Arbitres / Referees

Mathieu Blanchette (McGill)
Henner Brinkmann (UdeM)
Nadia El-Mabrouk (UdeM)
Francois Major (UdeM)
Vladimir Makarenkov, UQAM
Marcel Turcotte (Ottawa)

Renseignements généraux / General information

Accueil / Registration

L'accueil des participants se fera au Hall d'honneur du pavillon Roger-Gaudry (anciennement Pavillon principal), le mercredi 2 novembre dès 8 h 30. Les insignes d'identification vous seront remis à ce moment.

The registration office is located in the Honor Hall in the Roger-Gaudry Building (previously Main Building). Your identification badge will be available from 8:30 am, November 2.

Pauses santé et cocktail / Coffee breaks and cocktail

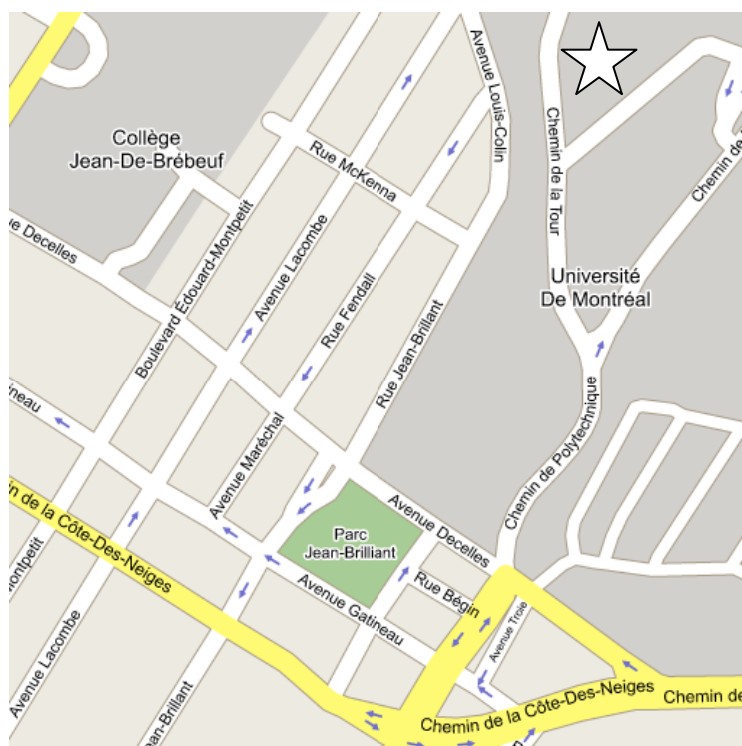
Les pauses santé et le cocktail seront servis dans le Hall d'honneur.

Coffee breaks and the cocktail party will be served in the Honor Hall.

Restaurants

Pour les repas du midi, de nombreux restaurants situés sur le Chemin de la Côte-des-Neiges ainsi que sur l'Avenue Lacombe sauront satisfaire tous les goûts.

For lunch, many restaurants are located on Chemin de la Côte-des-Neiges or Lacombe Avenue.



Horaire du mercredi 2 novembre 2005

- 9 h 15 Ouverture du colloque
Luc Vinet, recteur
- 9 h 30 Conférence : « Predictions of Protein-Protein Interactions »
Shoshana Wodak, Ph.D. - University of Toronto
- 10 h 30 OD1 : Data Rich But Information Poor : Applying Data Mining for Protein
Function Annotation
Sivakumar Kannan (UdeM)
- 11 h 00 Pause-santé
- 11 h 30 OD2 : Ontology-driven protein interaction data mining
Arash Shaban-Nejad (Concordia)
- 12 h 00 OD3 : L'analyse phylogénétique des protéases de la famille S2
Éric Simard (INRS-IAF)
- 12 h 30 Dîner
- 14 h 00 AM1 : *Pablo Zentilli (UQAM)*
AM2 : Trier par transpositions *Maxime Benoit-Gagné (UdeM)*
AM3 : Interaction gène-gène dans la susceptibilité à la leucémie de
l'enfant : approches statistiques et informatiques *Pascal St-Onge (UdeM)*
AM4 : *Fabrice Baro (UdeM)*
AM5 : *Guy Larochelle (UdeM)*
AM6: *Louis-Philippe Lavoie (UdeM)*
AM7 : RNA triloop characterization – RNA folding problem
Véronique Lisi (UdeM)
- 15 h 30 Pause-santé
- 16 h 00 OD4 : Phylogenomics - *Olivier Jeffroy (UdeM)*
- 16 h 30 OD5 : Evaluating the validity of composite taxa in phylogenic analysis
Véronique Campbell (UdeM)

OM =	Oral M.Sc.	AM =	Affiche M.Sc. Poster
OD =	Oral Ph.D.	AD =	Affiche Ph.D. Poster
OP =	Oral Postdoc		

Horaire du jeudi 3 novembre 2005

- 9 h 15 Mot de bienvenue : *Jacques Turgeon, vice-recteur à la recherche*
- 9 h 30 OD6 : Detecting and characterizing recombination - *Trevor Bruen (McGill)*
- 10 h 00 OD7 : Combining genes in phylogenetic analysis - *Rachel Bevan (McGill)*
- 10 h 30 OD8 : Les profils de substitution comme marqueur d'un changement fonctionnel - *Béatrice Roure (UdeM)*
- 11 h 00 Pause-santé
- 11 h 30 OD9 : Allelic variation for reconstructing the evolutionary history
Simon Joly (UdeM)
- 12 h 00 OD10 : Continuous model for evolutionary rate variation
Thomas Lepage (McGill)
- 12 h 30 OD11 : L'inférence phylogénétique avec des données manquantes
Abdoulaye Baniré Diallo (McGill)
- 13 h 00 Dîner
- 14 h 30 AD1 : Application of discrete event supervisory control in cancer treatment
Shausheen Zahirazami (Concordia)
AD2 : *Mathieu Lajoie (UdeM)*
AD3 : In silico identification of nucleus-encoded mitochondrial proteins using EST data - *Yaoqing Shen (UdeM)*
AD4 : Site-interdependant phylogenetic models - *Nicolas Rodrigue (UdeM)*
- 15 h 30 Pause-santé
- 16 h 00 AD5 : Caractérisation des interactions entre l'hémoglobine tronquée trHbN et le monoxyde d'azote par dynamique moléculaire - *Richard Daigle (Laval)*
AD6 : *Claudia Kleinman (UdeM)*
AD7 : Viral RNA : from sequence to structure - *Emmanuelle Permal (UdeM)*
AD8 : Prédire la structure 3D des protéines : Modélisation moléculaire par homologie et application à l'étude structurale d'un complexe protéine/acide nucléique - *Charlotte Habegger-Polomat (Laval)*
- 17 h 00 Clôture

Horaire du vendredi 4 novembre 2005

- 9 h 15 Mot de bienvenue : *Pierre Boyle, vice-doyen aux études supérieures et à la recherche (Faculté de Médecine)*
- 9 h 30
- 10 h 00 OM1 : Evolution of C2H2-Zinc finger genes in the Human Genome
Hamsa Tadepally (UdeM)
- 10 h 30 OM2 : RNA Triloop characterization – RNA folding problem
Véronique Lisi (UdeM)
- 11 h 00 Pause-santé
- 11 h 30 OM3 : Tetraloop classification : some observations
Nicolas Saint-Onge (UdeM)
- 12 h 00 OM4 : Interface web interactive pour le calcul et la visualisation d'arbres et de réseaux phylogéniques
Alpha Boubacar Diallo (UQAM)
- 12 h 30 Dîner
- 14 h 00 Large-scale phylogenetic reconstruction and the Tree of Life
Bernard Moret, Ph.D. (University of New Mexico)
- 15 h 00 Pause-santé
- 15 h 30 OP1 : Vers la modélisation du réseau des processus cellulaires
Vincent Devloo (UdeM)
- 16 h 00 OP2 : Effects of compositional heterogeneity on phylogenomic analyses
Denis Baurain (ULg-UQ-UdeM)
- 16 h 30 Remise des prix (4) et clôture : Jacques Bélair et Gertraud Burger
- 17 h 00 Cocktail

CONFÉRENCES

Bernard Moret
Department of Computer Science
University of New Mexico

Large-scale phylogenetic reconstruction and the Tree of Life

In this talk, I will review current computational activities aimed at reconstructing the Tree of Life, the evolutionary history of all living organisms (an estimated 10-100 million). Researchers and funding agencies worldwide have put renewed emphasis on the establishment of evolutionary relationships among living species; such relationships are fundamental to research in medicine, drug design, agriculture, ecology, and many other areas.

The CIPRES (Cyber Infrastructure for Phylogenetic Research), which I direct, was founded to develop the informatics tools required to attempt a reconstruction of the Tree of Life. I will sketch the goal and current achievements of CIPRES, comment on future needs, and relate its work to that of other research efforts in phylogeny. I will then focus on algorithmic techniques for large-scale reconstruction, both in terms of speed and in terms of accuracy. In the process, I will discuss novel divide-and-conquer approaches broadly applicable across all optimization criteria and types of data and also review progress in using one specific type of data, gene rearrangements.

Shoshana J. Wodak
Department of Biochemistry and Structural Biology
Department of Medical Genetics
University of Toronto

Predictions of Protein-Protein Interactions

With accrued interest in protein-protein interactions, the prediction of these interactions from sequence and structure information has become a booming activity. CAPRI (Critical Assessment of Predicted Interactions) the community-wide experiment for assessing blind predictions of protein-protein complexes is playing an important role in fostering progress in protein-protein docking algorithms – the procedures for predicting the 3D structure of protein complexes from those of the individual components-. In parallel, novel methods are being developed for predicting regions of a protein that are likely to interact, or for characterizing putative intermolecular contacts from sequence and structural data. Together with docking these methods provide an integrated computational approach that should be a valuable complement to the genome scale experimental studies of protein-protein interactions.

References

Wodak SJ, Mendez R. Prediction of protein-protein interactions: the CAPRI experiment, its evaluation and implications, *Curr Opin Struct Biol.* 2004 Apr;14(2):242-9

Mendez R, Leplae R, Lensink MF, Wodak SJ., Assessment of CAPRI predictions in rounds 3-5 shows progress in docking procedures., *Proteins.* 2005 Aug 1;60(2):150-69

PRÉSENTATIONS ORALES / ORAL PRESENTATIONS

OM1 : Hamsa Tadepally (UdeM)

Evolution of C₂H₂-Zinc finger genes in the Human Genome

Zinc finger genes of the C₂H₂ type represent the largest gene family of potential transcription factors in the Human genome.

To understand their evolutionary history, we searched for all Human C₂H₂ zinc finger genes. The identified 693 genes were found to be present in many clusters (84) spread out on different chromosomes. Our on-going study of their spatial and phylogenetic relationships within and between clusters will constitute a first step to better understand the organization and evolution of these genes and the mechanisms favouring their expansion in mammalian genomes.

OM2 : Véronique Lisi (UdeM)

RNA triloop characterization - RNA folding problem

Single strand RNA can adopt many conformations, each of which having a different impact on the final function of the molecule. Much research has been done recently to discover new RNA motifs that would have a biological meaning. However, the discovery of those new motifs can only express its full potential if it is combined with a formal, detailed classification based on the three-dimensional structure. We provide here the basis towards the development of a common vocabulary around RNA using the triloop motif as an example.

OM3 : Nicolas Saint-Onge (UdeM)

Tetraloop classification : some observations

Tetraloops motifs have been retrieved and analysed from a PDB data set containing both low resolution structures and virus structures. We have then classified all tetraloops according to different structure criteria: loop sequence, adjacent stacks found in the loop, stem sequence and RMSD (Root Mean Square Deviation). The classification aims to discover interesting structural and functional properties within the different groups so that we could eventually have a greater knowledge for predicting RNA folding.

OM4 : Alpha Boubacar Diallo (UQAM)

Interface web interactive pour le calcul et la visualisation d'arbres et de réseaux phylogénétiques

Nous avons développé une interface web regroupant un ensemble d'outils bioinformatiques en vue du calcul et de la visualisation d'arbres et de réseaux phylogénétiques. Trois outils implantés effectuent le prétraitement des données : l'alignement de séquences, la génération de matrices de distance à partir de séquences et enfin la génération de matrices d'arbres à partir d'un ensemble de lignées d'espèces. Trois nouveaux algorithmes pour l'inférence et la visualisation d'arbres et des réseaux phylogénétiques ont été conçus et implantés. Ces algorithmes peuvent être appliqués aux données phylogénétiques complètes ou partielles (e.g. matrices de distances incomplètes, séquences d'ADN avec des bases manquantes). Les réseaux phylogénétiques servent à détecter et à visualiser des transferts horizontaux de gènes, de même que les processus d'hybridation et d'homoplasie.

OD1 : Sivakumar Kannan (UdeM)

Data Rich But Information Poor: Applying Data Mining for Protein Function Annotation

Assigning function to hypothetical proteins is a classical problem in functional genomics. We developed an analysis procedure to identify protein function using predictive data mining. The goal is to detect hidden signatures and patterns in the integrated biological data, and to employ this new knowledge for deciphering genomic data at a large scale.

OD2 : Arash Shaban-Nejad (Concordia)

Ontology-driven protein interaction data mining

Due to diversity and complexity of applications and relevant queries in life science, it is unlikely that all of the biological knowledge can be represented with the traditional data management approaches. We describe how an ontology-driven knowledge can support and improve protein interaction data mining tasks for knowledge discovery in the area of fungal biology by providing simplified access to units of intersecting information from different biological databases and existing bio-ontologies and utilizing a logical framework in answering questions from biologists.

OD3 : Éric Simard (INRS-IAF)

L'analyse phylogénétique des protéases de la famille S2

La forte relation entre l'activité des protéases et l'adaptation écologique des organismes en fait un groupe d'un grand intérêt pour l'appréciation des processus évolutifs. Les travaux réalisés ont porté sur une nouvelle enzyme de la famille des chymotrypsines (SLP25). L'analyse phylogénétique de cette enzymes et des protéases de la famille S2 a permis de proposer la formation d'une sous-famille et de démontrer que SLP25 est en fait une preuve qu'un échange latéral d'un gène d'un procaryote à un eucaryote. Ce groupe d'enzyme a utilisé plusieurs voix évolutives dont la duplication, la recombinaison et le transfert latéral.

OD4 : Olivier Jeffroy (UdeM)

Phylogenomics

Phylogenomics, the use of large amounts of genomic data to build phylogenetic trees, was expected to end these incongruences by providing highly supported and correct trees. Using a dataset of 106 genes from 14 yeasts, we show that the use of different tree reconstruction methods using either nucleotide or amino acid sequences leads to highly supported, but nevertheless, mutually incongruent phylogenies.

OD5 : Véronique Campbell (UdeM)

Evaluating the validity of composite taxa in phylogenetic analysis

En analyse phylogénétique, l'approche des taxons chimères combine les séquences d'ADN disponibles à l'intérieur d'un groupe monophylétique, pour ne former qu'un seul taxon dont les séquences sont complètes. Ainsi un nombre maximal de caractères peuvent être utilisés. À l'aide de simulations, nous explorons la validité de cette approche en la comparant à l'analyse de matrices incomplètes. Une application à la phylogénie des mammifères sera également présentée.

OD6 : Trevor Bruen (McGill)

Detecting and characterizing recombination

Recombination is a powerful evolutionary force that merges historically distinct genotypes. But the extent of recombination within many organisms is unknown, and even determining its presence within a set of homologous sequences is a difficult question. I will overview some recent results related to the problem of detecting and characterizing recombination.

OD7 : Rachel Bevan (McGill)

Combining genes in phylogenetic analysis

Combining genes in phylogenetic analysis is becoming more common, especially since more data often leads to better support for the analysis. However, how best to combine such data is an open question, especially since different genes are often under different selective pressures. The importance of accounting for the "rate" of evolution of a gene in multi-gene analysis has been noted. In this talk I will discuss two approaches to this problem: Using gene rates as branch length multipliers and integrating out over all possible gene rates for a given gene.

OD8 : Béatrice Roure (UdeM)

Les profils de substitution comme marqueur d'un changement fonctionnel

Comme la duplication de gènes est couramment utilisée par la nature pour créer de nouvelles fonctions et que la puissance fournie par l'évolution a été sous-estimée, notre but est de déterminer des positions impliquées dans un changement fonctionnel en étudiant les propriétés évolutives d'une même position chez 2 paralogues. Par comparaison des profils évolutifs définis avec le modèle CAT, nous avons démontré que ces profils sont biologiquement significatifs. Comme la fréquence de profils différents en une position est plus grande entre paires de paralogues qu'entre paires d'orthologues, ce critère fournit une méthode précise pour trouver des positions d'intérêt.

OD9 : Simon Joly (UdeM)

Allelic variation for reconstructing the evolutionary history

Although allelic variation holds information regarding the relationships of organisms, it is rarely included in phylogenetic analyses. Here we propose a distance-based method that incorporates allelic variation for reconstructing the phylogeny of organisms from one or more genes. As an example, the method is used to reconstruct the phylogeny of North American roses.

OD10 : Thomas Lepage (McGill)

Continuous model for evolutionary rate variation

We propose a continuous model for evolutionary rate variation across sites and over the phylogenetic tree and derive exact transition probabilities of substitutions under this model. Our model is based on the CIR process, a diffusion widely used in financial applications. This model directly extends the standard gamma distributed rates across site model, with one additional parameter governing changes in rate down the tree. Markov Chain Monte Carlo methods can be employed under the CIR model for computing the likelihood of trees and to test heterotachy. Some applications are provided.

OD11 : Abdoulaye Baniré Diallo (McGill)

L'inférence phylogénétique avec des données manquantes

PEMV est une nouvelle façon de considérer les bases manquantes lors de l'inférence phylogénétique. Cette méthode d'estimation des valeurs manquantes dans les séquences nucléotidiques permet d'améliorer la qualité d'inférence phylogénétique par rapport à plusieurs méthodes existantes. Elle donne également des résultats prometteurs pour l'inférence de phylogénie à partir de séquences multigéniques. Je termine en présentant une adaptation de la méthode au programme PHYML.

OP1 : Vincent Devloo (UdeM)

Vers la modélisation du réseau des processus cellulaires

Lors de la modélisation d'un réseau génétique, on cherche un modèle qui tienne compte du réseau d'interactions et des états stables du système étudié. Nous avons recherché deux types de sous-réseaux particuliers: le réseau nécessaire et les réseaux minimaux. Un réseau nécessaire est le réseau où on a conservé seulement les interactions nécessaires à l'obtention d'un modèle, considérant les états stables du système. Un réseau minimal est un modèle valide mais aucune interaction ne peut être enlevée.

OP2 : Denis Baurain (ULg – UQ - UdeM)

Effects of compositional heterogeneity on phylogenomic analyses

In a recent study, Beiko et al. (2005) have investigated the extent of lateral gene transfer in prokaryotes by performing a rigorous phylogenetic analysis of over 220,000 proteins from 144 genomes. While the inferred relationships suggest a largely vertical inheritance pattern, discordance between individual trees and the reference supertree pinpoints 'highways' of gene sharing between lineages. We propose to use the four picocyanobacteria in this dataset as a test case to determine whether such discordance could arise from methodological issues (e.g. compositional heterogeneity).

AFFICHES / POSTERS

AM1 : Pablo Zentilli (UQAM)

The hit selection procedure in HTS screens requires routines for quality control and data correction. We present new methods and corresponding software, called HTS Corrector, intended for statistical analysis and visualization of HTS data. These methods are based on the statistical analysis of signal variation within plates of an assay. They allow researchers to analyze experimental HTS assays, detecting and characterizing systematic errors. A k-means partitioning procedure divides original dataset on several subsets with similar systematic patterns. Two algorithms for data correction are implemented in HTS Corrector. The sigma distribution analysis allows users to adjust sigma level for hit selection for each particular well. HTS Corrector displays color plate maps, tables, and histograms. The executable file for Windows and the detailed description of the software are freely available on our website: <http://www.labunix.uqam.ca/~makarenv/hts.html>.

AM2 : Maxime Benoit-Gagné (UdeM)

Trier par transpositions

Un algorithme approximatif pour estimer la distance entre 2 génomes en se basant sur un ordre des gènes et l'opération de réarrangement génomique nommée transposition. Celle-ci consiste à "couper" un bloc de gènes et à le "coller" ailleurs dans le génome.

AM3 : Pascal St-Onge (UdeM)

Interactions gène-gène dans la susceptibilité à la leucémie de l'enfant: approches statistiques et informatiques

La leucémie lymphoblastique aiguë (LLA) est le cancer le plus fréquent chez les enfants. Notre groupe de génétique du cancer a démontré l'importance jouée par la variabilité inter-individuelle dans la susceptibilité à la LLA. L'analyse simultanée de plusieurs loci, basée sur des modèles statistiques existants, permettra de vérifier la présence d'effets synergiques entre différentes formes de gènes impliqués dans le contrôle du cycle cellulaire, la réparation de l'ADN et le métabolisme des carcinogènes.

AM4 : Fabrice Baro (UdeM)

Les transferts horizontaux de gènes (HGT) consistent en l'échange de matériel génétique entre des individus d'espèces différentes. Lors de la reconstruction d'un arbre phylogénétique avec des espèces impliquées ayant subi des HGT, ces dernières apparaissent plus proches qu'en réalité. Nous essayons de quantifier et caractériser les HGT, notre hypothèse étant que les gènes les plus répandus sont moins transférés. Chez les bactéries dont le génome est séquencé, nous sélectionnons les gènes "orthologues", nous contruisons la phylogénie de chaque groupe d'orthologues, puis nous déterminons s'il y a eu transfert en la comparant avec la phylogénie supposée des organismes.

AM5 : Guy Laroche (UdeM)

Dans le but de pouvoir essayer de prédire l'hétérodimérisation des récepteurs couplés aux protéines G (RCPG), certaines règles ont été déterminées pour que 2 récepteurs puissent avoir ce type d'interaction : ils doivent être co-exprimés dans les mêmes tissus au même moment et posséder des régions régulatrices similaires. Par conséquent, le but du présent projet est d'identifier des régions conservées qui pourrait servir à faire une seconde classification des RCPG selon laquelle ces récepteurs seraient classés en fonction de leur susceptibilité d'hétérodimériser avec les autres récepteurs.

AM6 : Louis-Philippe Lavoie (UdeM)

ARN structural study through cyclic decomposition; Using a cyclic description method developed in our lab, ARN structures can be compared to search for global and/or local similarities or differences, perform genuine conformational alignment and classify molecules based on their tertiary structure.

AM7 : Véronique Lisi (UdeM)

RNA triloop characterization - RNA folding problem

Single strand RNA can adopt many conformations, each of which having a different impact on the final function of the molecule. Many research has been done recently to discover new RNA motifs that would have a biological meaning. However, the discovery of those new motifs can only express its full potential if it is combined with a formal, detailed classification based on the three-dimensional structure. We provide here the basis towards the development of a common vocabulary around RNA using the triloop motif as an example.

AD1 : Shauheen Zahirazami (Concordia)

Application of discrete event supervisory control in cancer treatment

The class of discrete event systems can be formally modeled using finite state machines. Cells of organisms can be viewed as systems driven by metabolic pathways which are regulated by feedback mechanisms and can also be modeled as DES. Gene expression and protein interactions are examples of different events which derive the state of the cell in a huge automata model.

The ultimate goal of this research is to come up with a DES model of cancer to develop early detection methods, accurate prognosis and effective therapies, by utilizing state estimation, computational simulation and failure recovery techniques.

AD2 : Mathieu Lajoie (UdeM)

Les conversions géniques jouent un rôle important dans l'évolution des populations eukaryotes en échangeant de courts segments d'ADN durant la méiose. Les expérimentations permettant d'étudier directement ce phénomène sont complexes et coûteuses, d'où l'attrait de développer des approches informatiques d'inférence qui utilisent les haplotypes. Nous présenterons une nouvelle approche qui pourrait permettre d'estimer les taux de conversion génique avec plus de précision.

AD3 : Yaoqing Shen (UdeM)

In silico identification of nucleus-encoded mitochondrial proteins using EST data

In this study we tested the performance of five available programs on predicting mitochondrial proteins from 433 jakobid ESTs. Only 30 to 60% of the ESTs are predicted correctly. The prediction results are analyzed to see why some ESTs are correctly predicted while others not. We then trained and tested one program with Arabidopsis ESTs. Again, only 20% ESTs are correctly predicted. Therefore, now we are working on a new method by which prediction of mitochondrial proteins using ESTs sequences should be more efficient.

AD4 : Nicolas Rodrigue (UdeM)

Site-interdependent phylogenetic models

The assumption of site-independence invoked in most phylogenetic studies is known to be biologically unsound. We have recently applied the sampling methods of Robinson et al. (Mol. Biol. Evol. 2003) to propose a model that explicitly accounts for site-interdependencies resulting from protein tertiary structure constraints (Rodrigue et al., Gene, 2005). Through two Bayesian techniques, we show that site-interdependence due to protein tertiary structure is always favored over standard models.

AD5 : Richard Daigle (Laval)

Caractérisation des interactions entre l'hémoglobine tronquée trHbN et le monoxyde d'azote par dynamique moléculaire

Mycobacterium tuberculosis (TB), causant la tuberculose, synthétise des enzymes appelées hémoglobines tronquées (trHbN) qui détoxifient le monoxyde d'azote (NO) produit par les macrophages pour contrer les défenses de l'hôte. Des études de dynamique moléculaire ont été menées dans notre laboratoire afin de caractériser la dynamique de l'interaction entre trHbN et le NO. Ces travaux constituent une percée majeure dans la compréhension des mécanismes de diffusion du NO au coeur de trHbN.

AD6 : Claudia Kleinman (UdeM)

We present a set of statistical potentials specifically devised to estimate the compatibility of a sequence to a particular protein structure. All the necessary parameters are optimized from a database of protein structures with a gradient descent procedure. We show the general behavior of the optimization process as well as the dependence on the protein representation. All together, the probabilistic framework proposed makes it possible to test a diversity of models and choose the best suited to the protein design problem.

AD7 : Emmanuelle Permal (UdeM)

Viral RNA: from sequence to structure

The integration of genomic and structural data can bring a powerful tool for studying RNA viruses.

This poster illustrates a way of performing analysis and modeling from a tertiary structure file using the stemloop D of coxsackievirus B3 (PDB 1RRF) as an example.

Three models of stemloop D are proposed : for Poliovirus, Human Rhinovirus 89 and Human Enterovirus C (based on sequence conservation within picornaviridae) and structural features of 1RFR are analysed.

AD8 : Charlotte Haberger-Polomat (Laval)

Prédire la structure 3D des protéines: Modélisation moléculaire par homologie et application à l'étude structurale d'un complexe protéine/acide nucléique

La modélisation moléculaire est un domaine d'application des outils de la bio-informatique en plein essor. Ici seront présentés les principes généraux de la modélisation moléculaire, les techniques utilisées et les logiciels disponibles.

Le programme de modélisation 3D par homologie Modeller, utilisé au laboratoire, sera détaillé, ainsi que nos résultats préliminaires obtenus avec cet outil pour notre protéine d'intérêt: la glutamyl-ARNt synthétase d'E. coli, en interaction avec son ARNt.

-Madore E, Lipman RS, Hou YM, Lapointe J. (2000) Evidence for unfolding of the single-stranded GCCA 3'-End of a tRNA on its aminoacyl-tRNA synthetase from a stacked helical to a foldback conformation. *Biochemistry*, 39(23):6791-8.

-Sali A. & Blundell TL. (1993) Comparative protein modelling by satisfaction of spatial restraints. *J. Mol. Biol.* 234, 779-815.