



La bioinformatique présentée par un doctorant

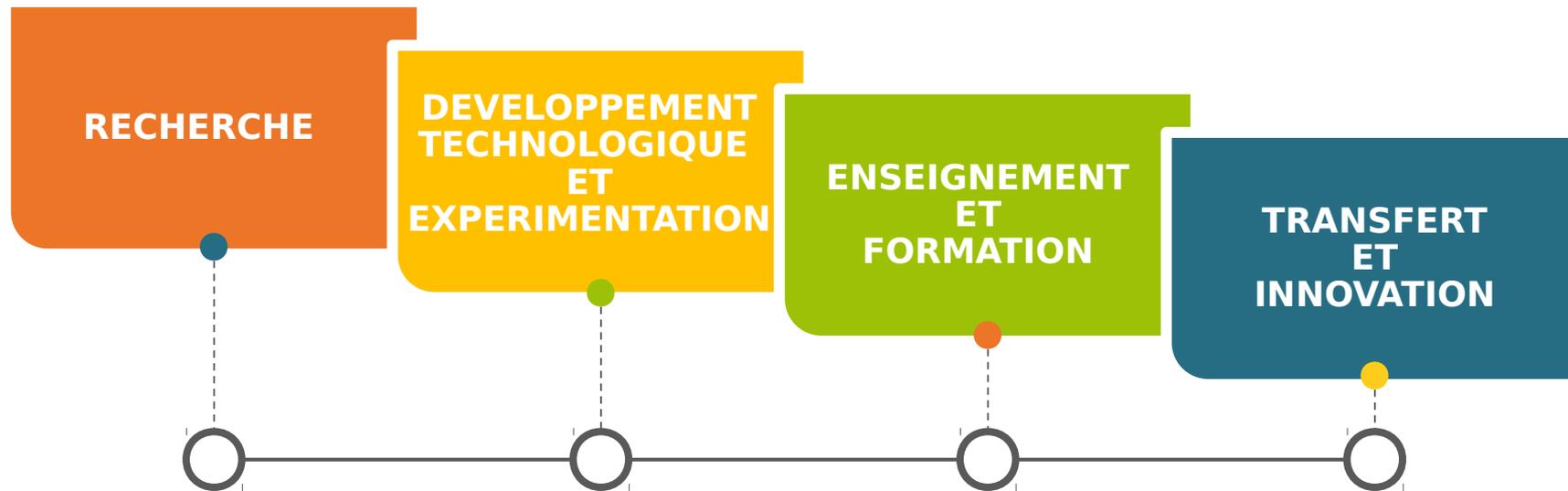
1

La recherche en France

La recherche en France

- 393.000 personnes dont 239.000 chercheurs
 - Recherche divisée en différents organismes :
 - C.N.R.S. (un peu de tout)
 - Inria (informatique et mathématiques)
 - Inra (agronomie)
 - Ifremer (mer)
 - Inserm (santé)
 - CNES (études spatiales)
 - etc
- 35 organismes de recherche sous tutelle du ministère de l'enseignement supérieur et de la recherche
 - 75 universités
 - Grandes écoles
 - etc

Inria : l'institut de recherche en sciences du numérique



Un institut public à caractère scientifique et technologique
sous la double tutelle du ministère de la Recherche
et du ministère de l'Industrie

Principaux domaines de recherche

- 1 Mathématiques appliquées, Calcul et Simulation
- 2 Algorithmique, Programmation, Logiciels et Architectures
- 3 Réseaux, Systèmes et Services, Calcul distribué
- 4 Perception, Cognition, Interaction
- 5 STIC pour les sciences de la vie et de l'environnement

8 centres de recherche Inria



Inria PARIS - Rocquencourt



Inria LILLE
Nord Europe



Inria NANCY
Grand Est



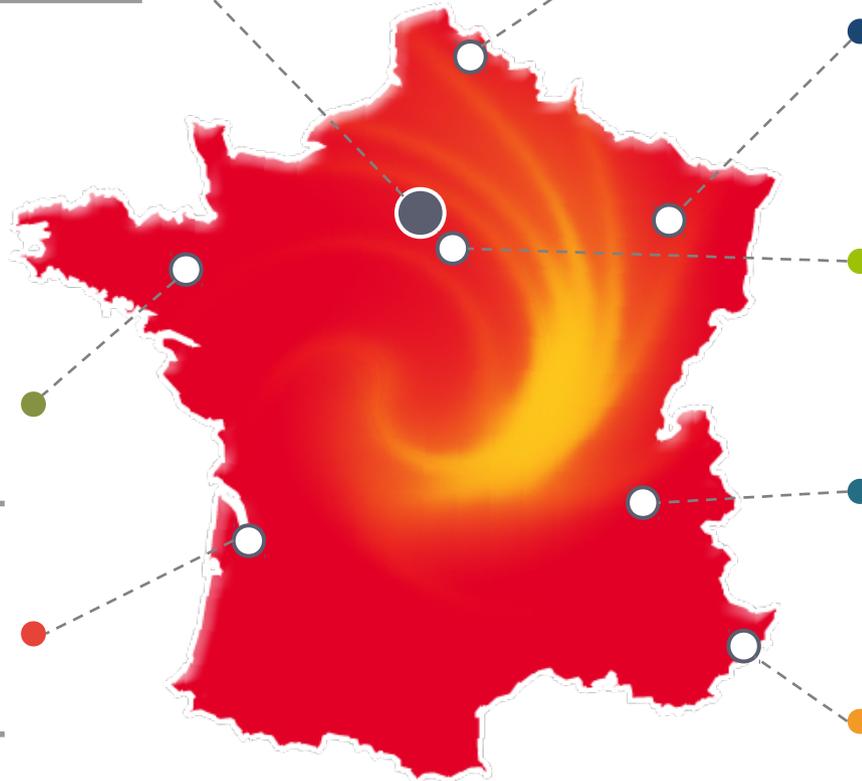
Inria SACLAY
Île-de-France



Inria GRENOBLE
Rhône-Alpes



Inria SOPHIA ANTIPOLIS
Méditerranée



Inria RENNES

Bretagne Atlantique



Inria BORDEAUX

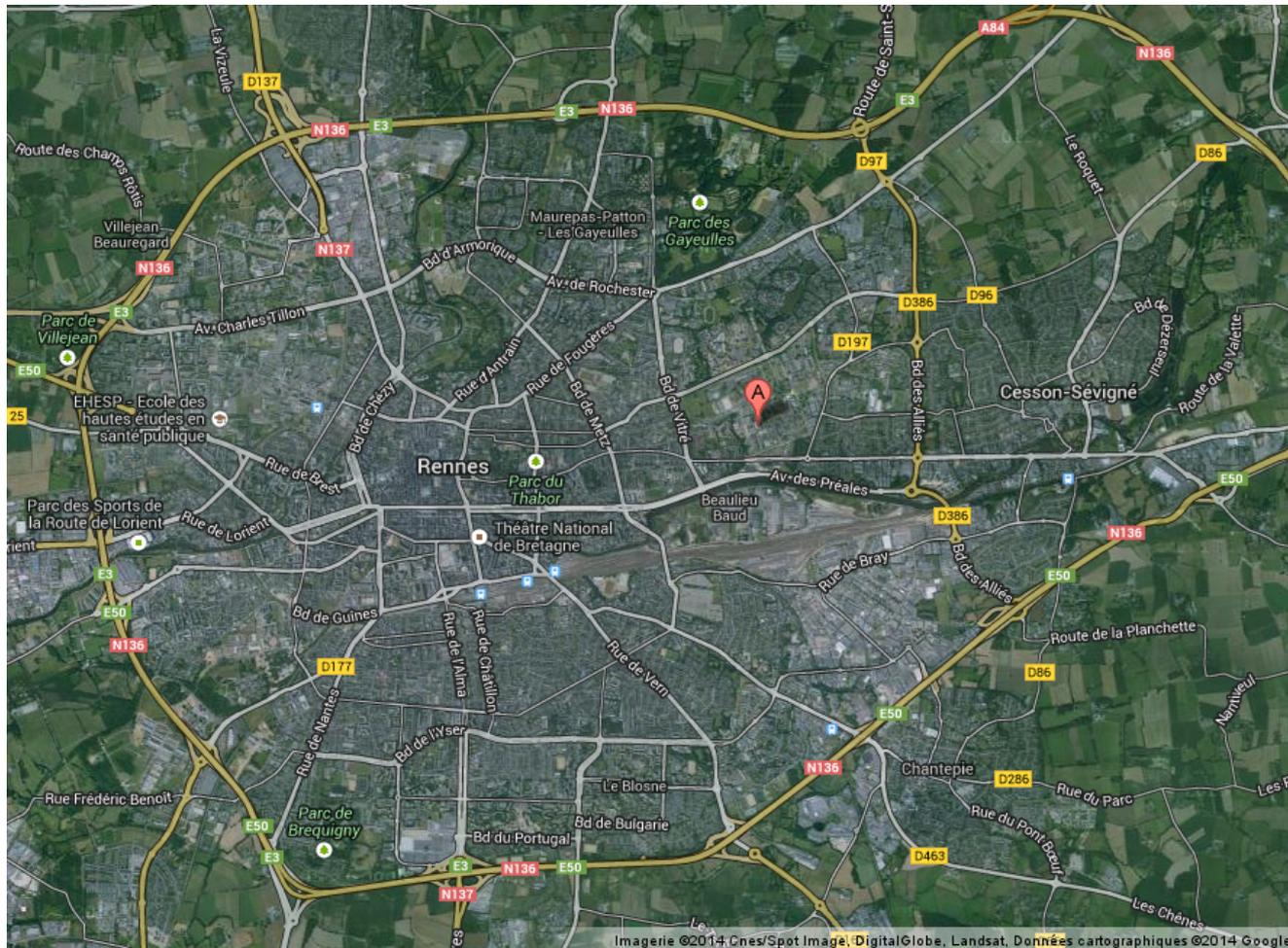
Sud-Ouest



Le centre Inria Rennes - Bretagne

Atlantique

- Situé sur le campus de Beaulieu



Le centre Inria Rennes - Bretagne Atlantique

- Situé sur le campus de Beaulieu



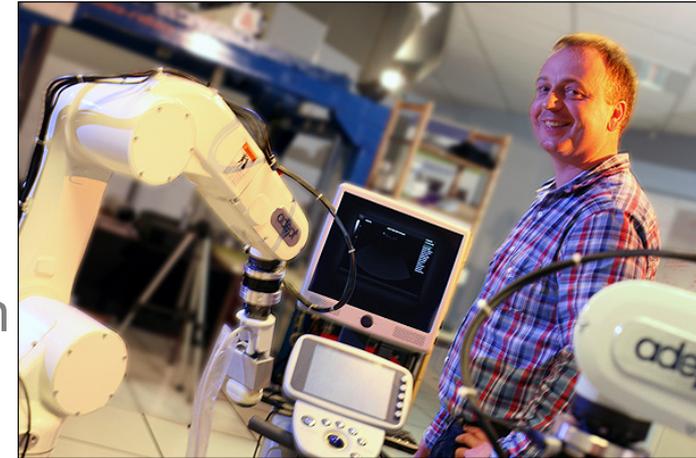
Le centre Inria Rennes - Bretagne Atlantique

- Qui y travaille ?
 - Environ 600 personnes
 - 400 scientifiques
 - 200 personnes dédiées au support et à l'accompagnement de la recherche
 - 47 nationalités représentées
 - 3 sites : Rennes, Nantes et Lannion



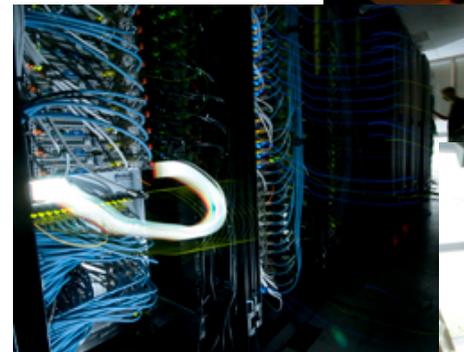
Le centre Inria Rennes - Bretagne Atlantique

- Qu'y faisons nous ?
 - Algorithmique, programmation, architecture
 - Mathématiques appliquées, calcul, simulation
 - Perception, cognition, interaction
 - Réseaux, calcul distribué
 - Santé, biologie, planète



Le centre Inria Rennes - Bretagne Atlantique

- Des plateformes technologiques
 - Multimédia
 - Salle immersia, une des plus grandes salles de réalité virtuelle au monde
 - Informatique
 - Grid 5000, Genouest
 - Robotique
 - Industrielle, mobile, médicale
 - Biomédical
 - Neurinfo



Sylvain Prigent, ma vie, mon œuvre

- Mon parcours, où comment j'ai atterri à Inria ?

- Juillet 2002 : Bac S, spécialité physique-chimie



- 2002-2003 puis 2003-2004 : deux échecs au concours d'entrée en pharmacie

- 2004-2005 puis 2005-2006 : idem mais en médecine



- 2006 : entrée à la fac en seconde année de biologie



- 2008 : obtention d'une licence de biologie et entrée en master de bio-informatique

- 2010 : obtention du master de bio-informatique

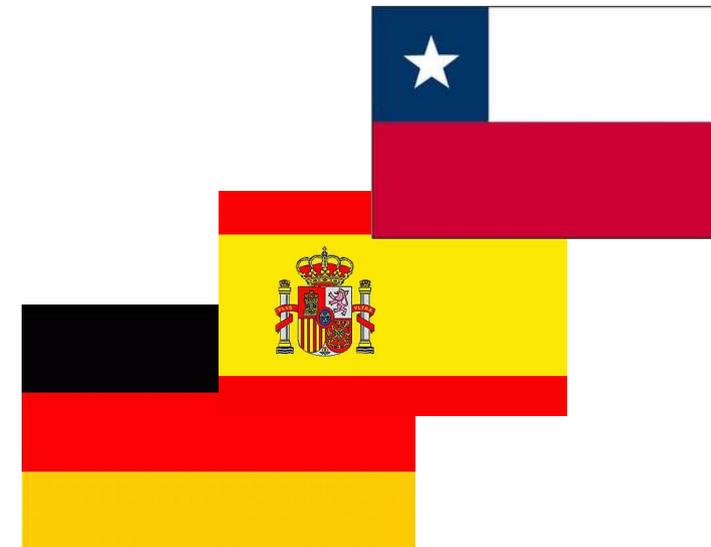


- 2011 : obtention d'un second master, en informatique, et début de ma thèse

Le doctorant, un chercheur pas comme les autres



- Étudiant (bac + 5)
- Activité de recherche à part entière
- Porte un projet de recherche sur 3 ans (en France)
 - Relation conflictuelle avec la thèse
- Est rémunéré, cotise pour la sécu, la retraite, etc
- Voyage beaucoup
- Termine son projet de recherche par la rédaction et la soutenance d'une thèse
- Devient alors un docteur

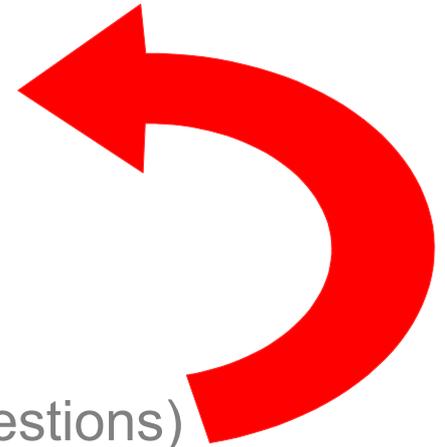


La vie du doctorant au jour le jour

- Possède un(e) directeur(trice) de thèse
- Est inscrit à une école doctorale qui suit son parcours
- Fait de la recherche, beaucoup de recherche !!
- Participe à de nombreux colloques, souvent à l'étranger
- Essaye de multiplier les collaborations, si possible internationales
- Donne des cours
- Essaye de conserver une « vraie » vie en dehors de sa thèse

La vie dans la recherche en informatique au jour le jour

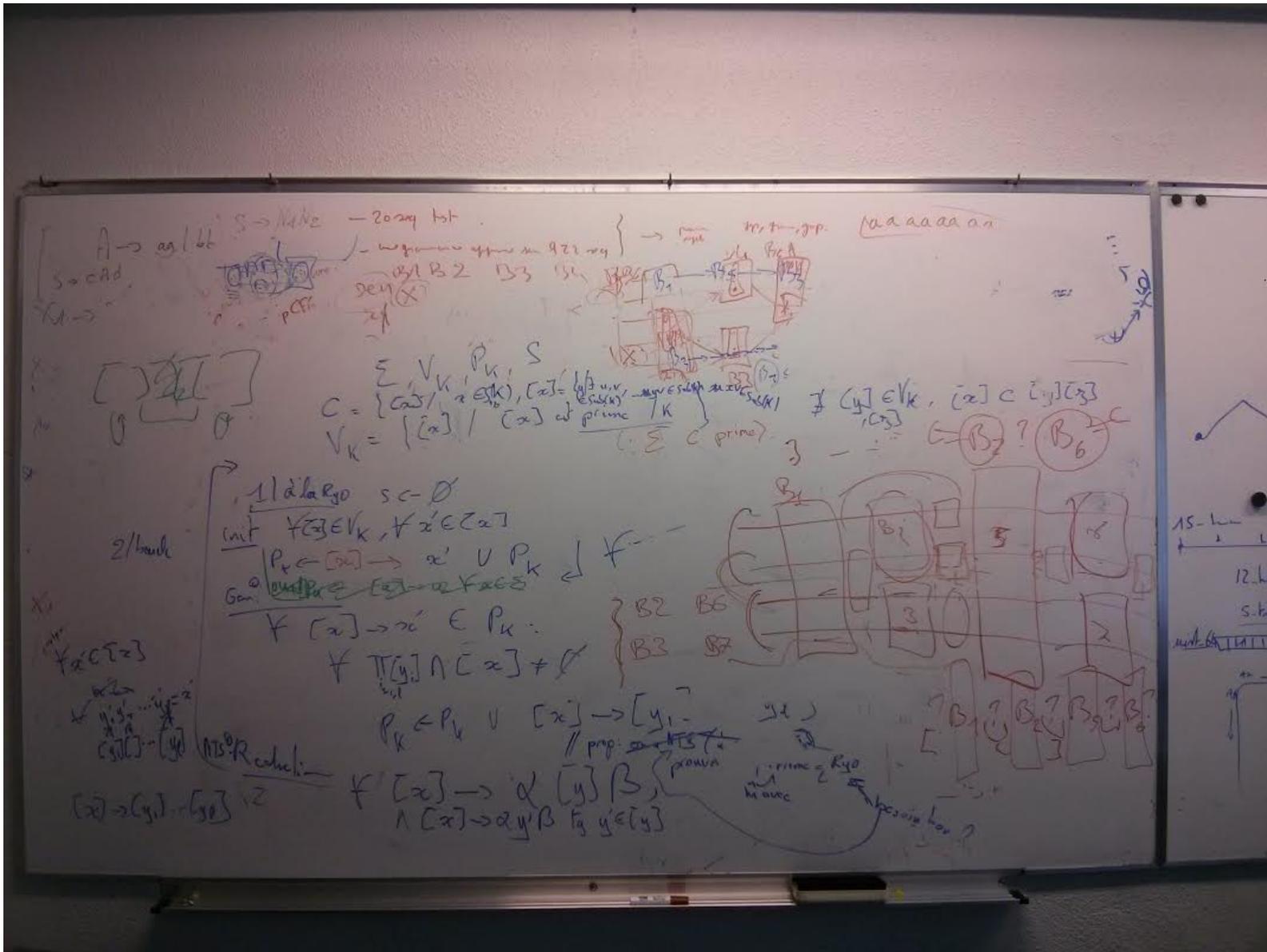
- Derrière notre ordinateur
 - Question
 - Invention d'algorithmes, de logiciels, de méthodes
 - Test de tout cela sur des exemples réels
 - Apport d'une réponse (ou souvent de nouvelles questions)
- Énormément d'échanges
 - Avec les chercheurs de notre équipe
 - Avec les chercheurs venant d'ailleurs
 - Avec tout le monde...



Et en dehors de la « vraie » recherche ?

- Communiquer !
 - Publications et exposés dans des conférences
 - Aller voir les autres chercheurs pour trouver de l'aide
 - Évaluer le travail des autres
 - Médiation scientifique
- Chercher de l'argent
- Encadrer des étudiants, enseigner
- Etc etc etc

Les clichés sur le chercheur



son

irs qui

Les possibilités de carrière offertes par la recherche

Professors

- « Le chercheur » n'existe pas...

Public

- Avec enseignement :
 - Maître de conférences
 - Professeur
- Sans enseignement
 - Ingénieur d'étude
 - Ingénieur de recherche
 - Chargé de recherche
 - Directeur de recherche



What my parents think I do



What my friends think I do



What my students think I do



What my spouse thinks I do



What my colleagues think I do



What I actually do



Privé

- Beaucoup plus flou
- Nombreuses possibilités de carrière

La recherche publique, ses bons et ses moins bons cotés

- Les points positifs
 - Travail passionnant
 - Voyages
 - Liberté presque totale
 - Beaucoup de rencontres
 - Sécurité de l'emploi
 - Renouvellement permanent du travail
- Les points négatifs
 - Bac + 8 minimum
 - Difficulté à obtenir un poste permanent
 - Salaire limité par rapport au privé
 - Travail passionnant...

2

La Bioinformatique

La bioinformatique, petite présentation

- Champ de recherche inter-disciplinaire
 - Biologie, informatique, mathématiques, physique
- Discipline très très récente :
 - Terme utilisé pour la première fois en 1970
- Accumulation de quantités phénoménales d'information par les biologistes
 - L'outil informatique devient indispensable pour les étudier



La bioinformatique

- Un petit film d'introduction : **Bioinformaticus**
 - Festival Sciences en Cour[t]s
 - Prochaine édition : 15 mai 2014 au diapason
 - 3 doctorants de notre équipe de recherche

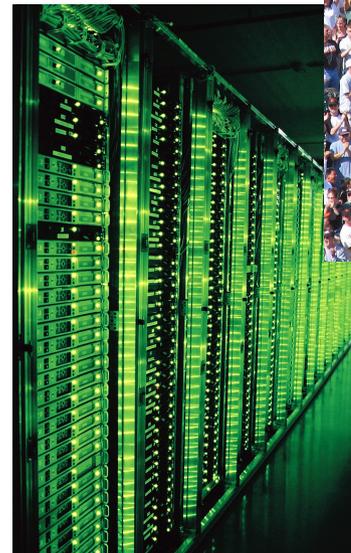
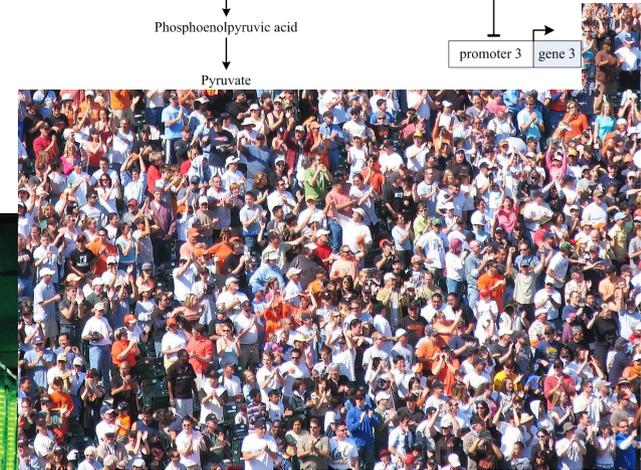
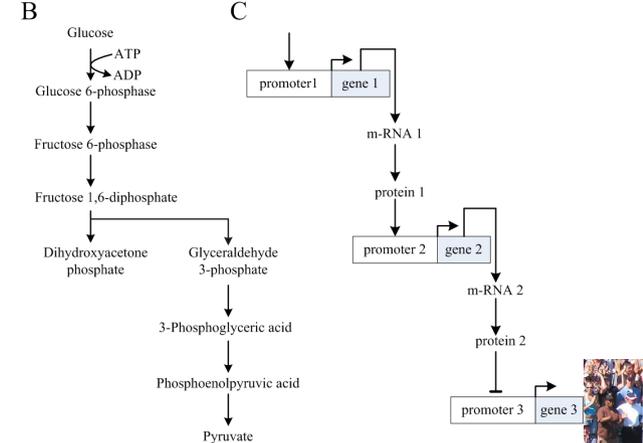
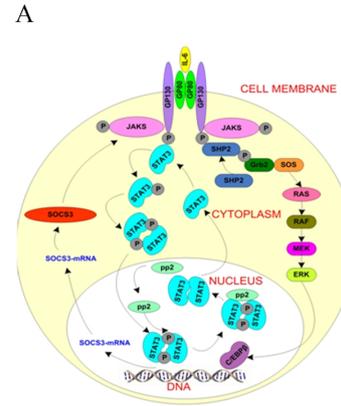


La bioinformatique à Inria Rennes



- 3 grands axes

- Modélisation du vivant
- Grosses masses de données
- Calcul intensif



Grosses masses de données

- Exemple de l'expédition Tara-océan
 - Echantillonnage des océans
 - Stockages de milliers de téra-octets de données sur place
 - Photos
 - ADN
 - Etc

Il faut réussir à
étudier tout ça !!!



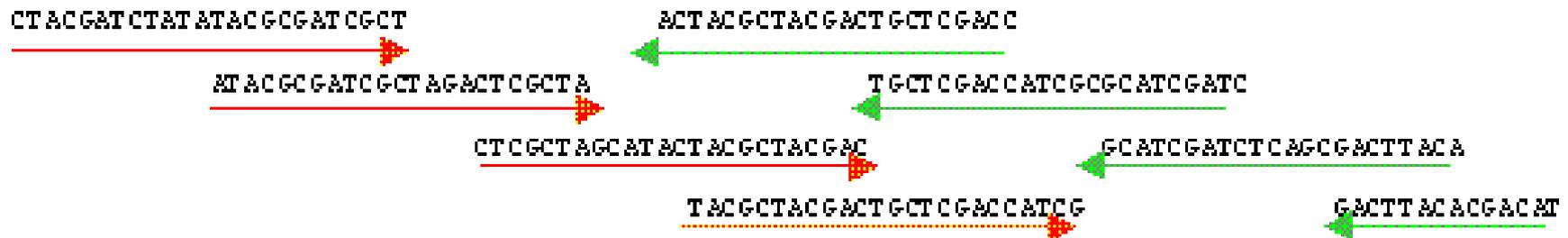
Calcul intensif



- La plateforme GenOuest
 - 276 cœurs et
~120 To de stockage au service de la biologie
- Le matériel ne fait pas tout !
 - Nécessité de travailler intelligemment
 - Une catapulte peut tuer une mouche. Mais est-ce vraiment indispensable ?
 - Gros travail d'optimisation pour accélérer la recherche et la rendre plus accessible à tous les laboratoires

Des domaines d'application connus

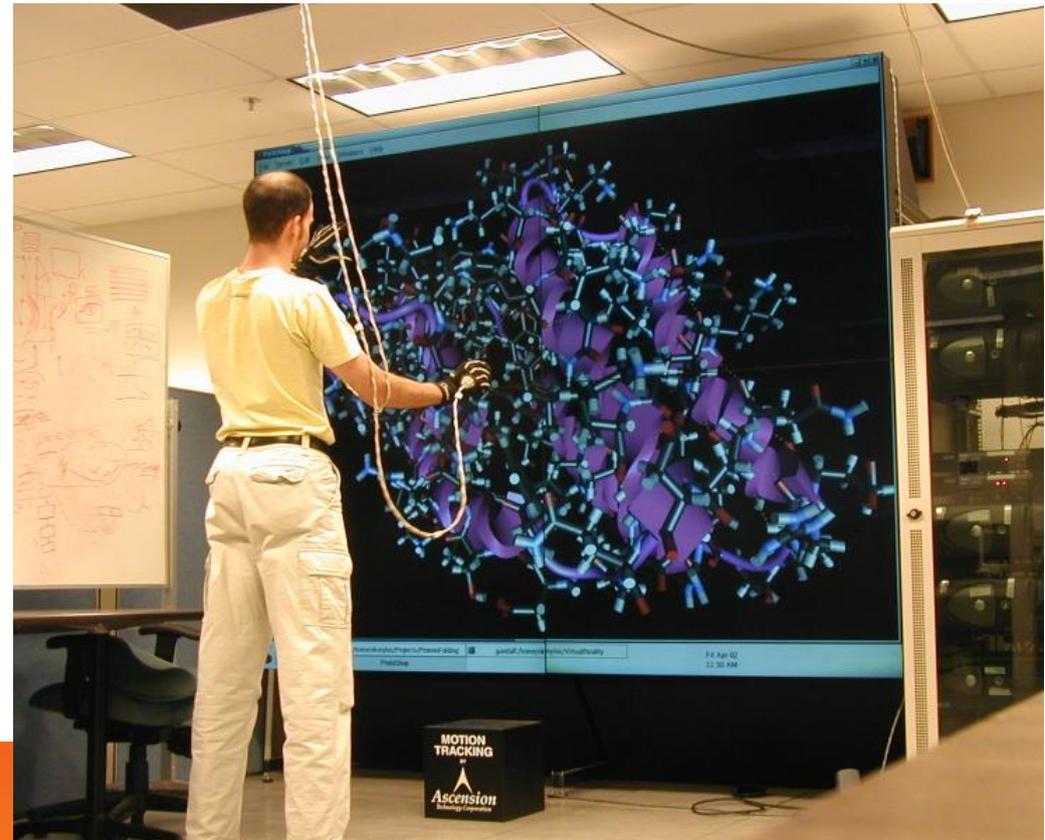
- Le séquençage et l'assemblage de l'ADN
 - Les « séquenceurs » nous donnent des petites séquences d'ADN
 - Des régions « chevauchantes » vont permettre de coller tout ces mots les uns à la suite des autres



- Assemblage de métagénomés ?

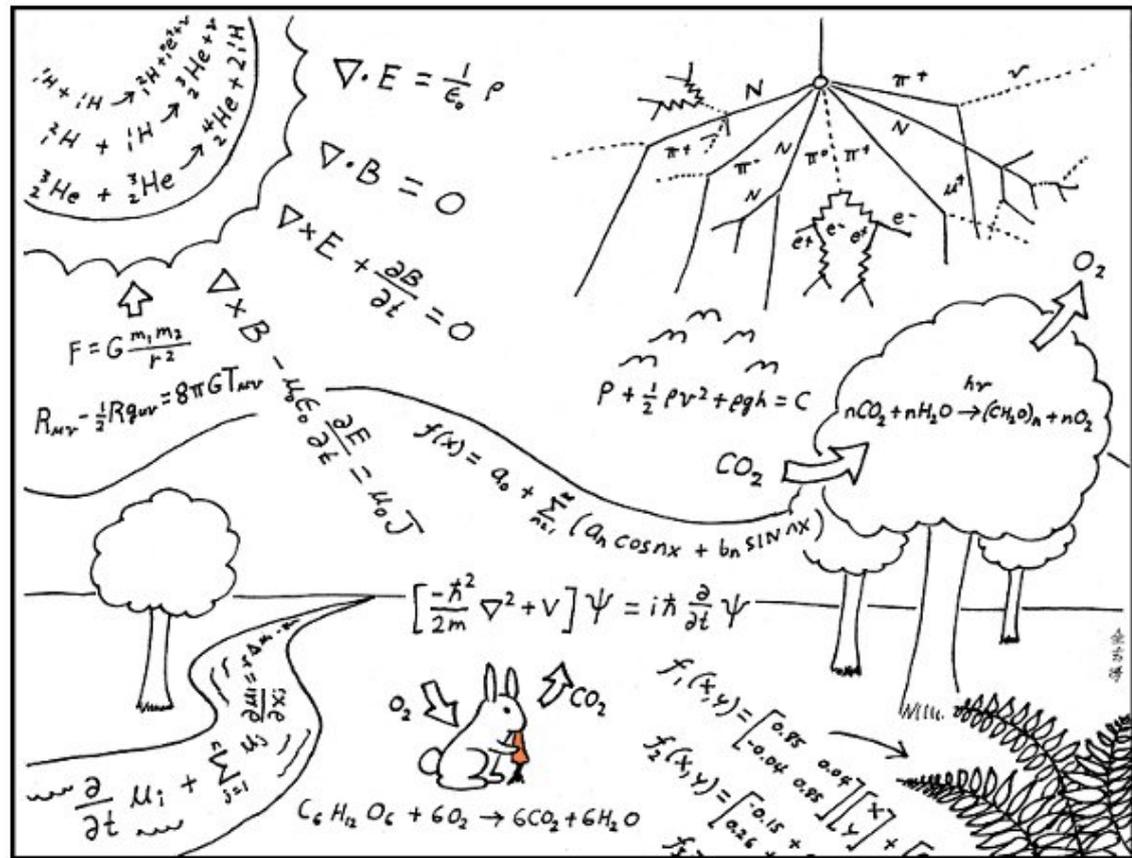
Des domaines d'application connus

- La modélisation en 3D de protéines
 - La « cristallographie » permet d'obtenir des structures tridimensionnelles des protéines
 - Cher, long, ne fonctionne pas toujours
 - Utiliser les connaissances précédentes pour accélérer tout ça
 - Utilisation de la réalité virtuelle, du retour haptique et de la bioinformatique dans une seule application



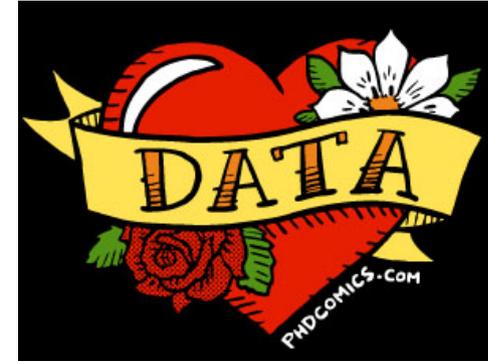
La modélisation

- Représentation schématique de la réalité
- Oubli volontaire (ou non...) de grandes parties de la réalité
 - Ex : carte routière
- Sert à répondre à une question précise !



This is how scientists see the world.

La modélisation de systèmes biologiques

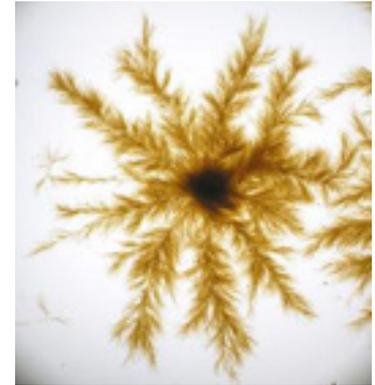


- Les biologistes possèdent :
 - Des données très précises (mais peu nombreuses)
 - Des données très nombreuses (mais peu précises)
- Le bio-informaticien fait le tri dans ces données pour apporter de nouvelles connaissances
 - Nouvelle utilisation de connaissances parfois vieilles de plusieurs dizaines d'années (voir siècles?)

La modélisation de systèmes biologiques appliquée à *Ectocarpus siliculosus*

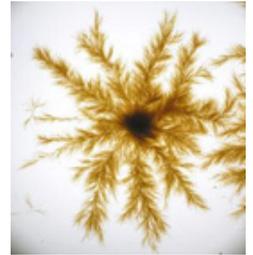
- *Ectocarpus*, une algue utile ?
 - Trop petite pour être utilisée en industrie
- *Ectocarpus*, une algue magnifique ?
- *Ectocarpus*, une algue étudiée ?
 - Partout dans le monde

- Mais pourquoi ???
 - C'est un organisme « modèle »



Ectocarpus, une algue modèle et top

- Petite, aisément cultivable en laboratoire
- Présente partout dans le monde (ou presque)



- Génome séquencé et annoté
- Très grande résistance aux variations de quantité de sel
- Semblable aux laminaires

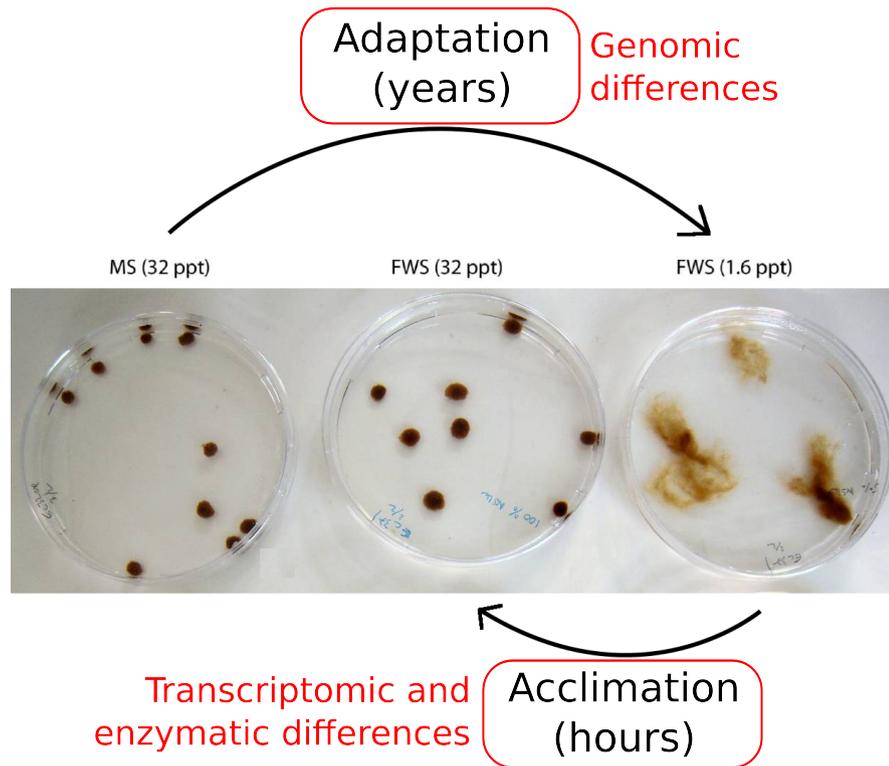


Station biologique de Roscoff

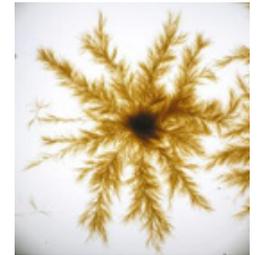
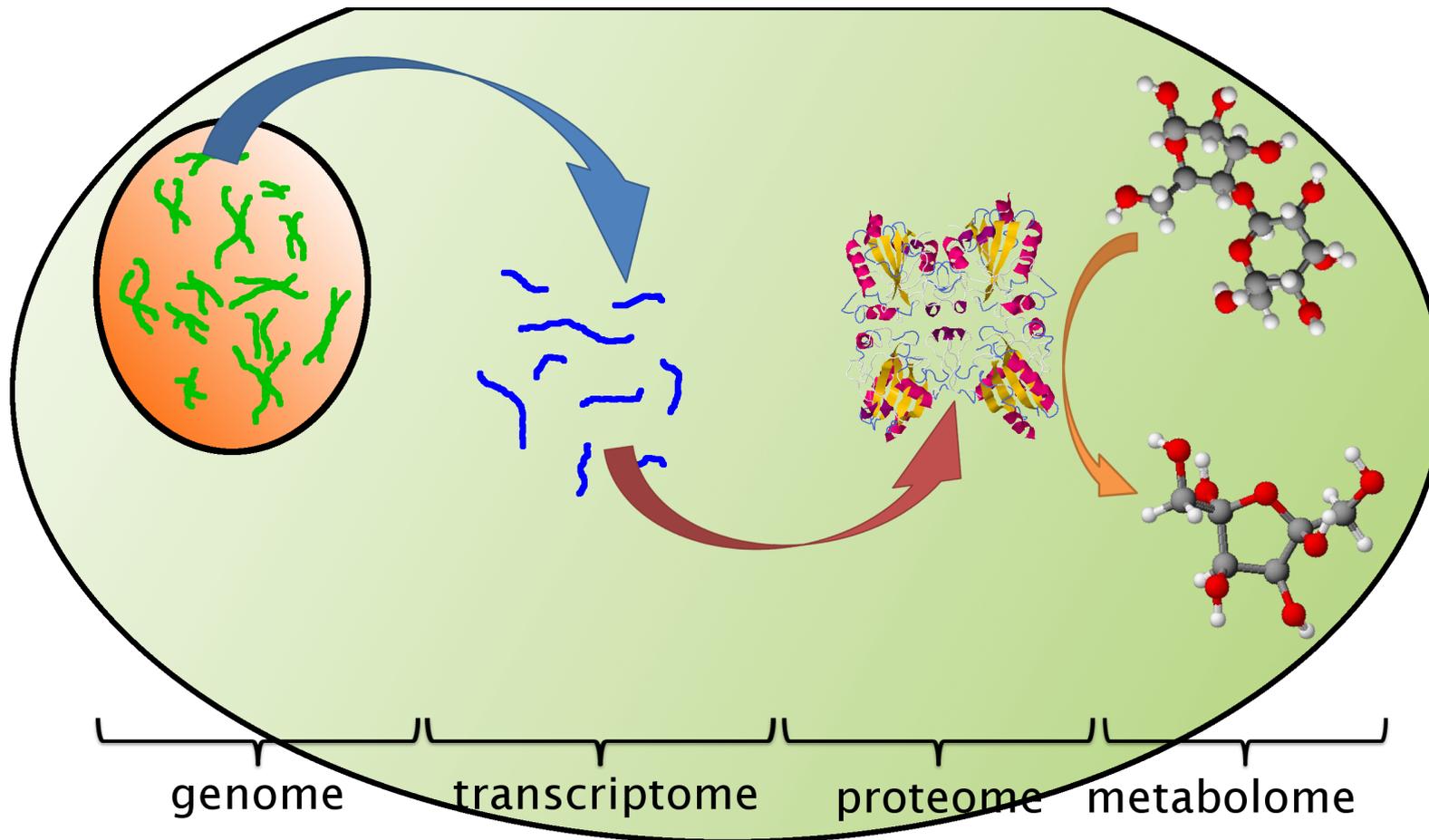


Ectocarpus, une algue « étrange »

- Une adaptation à l'eau douce très intéressante !

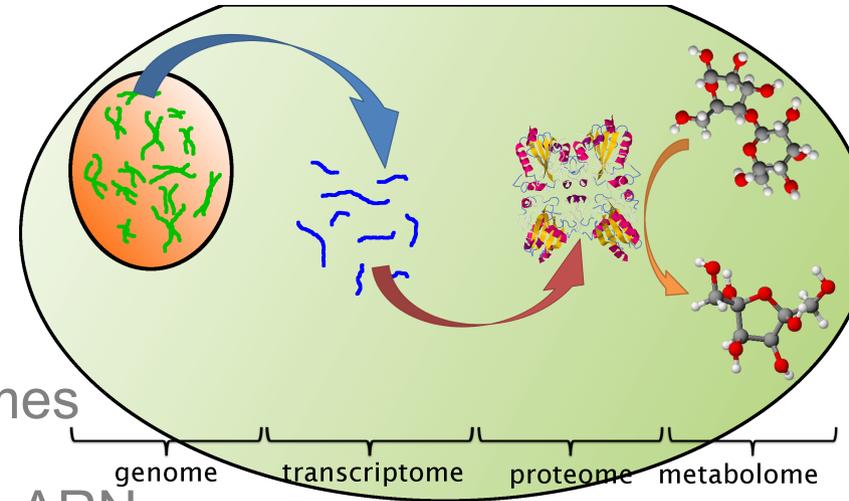


Où se situe cette forte tolérance au sel ?



Où se situe cette forte tolérance au sel ?

- Réponse rapide (marées)
 - Changements dans l'activité des enzymes
 - Changements dans la transcription des ARN
- Réponse lente (adaptation à l'eau douce)
 - Changements au niveau du génome



On souhaite étudier tout ces changements en même temps

Les réseaux métaboliques

- Modèles intégrant une grande majorité des données disponibles

Gène A
Gène B
Gène C



Protéine A
Protéine B
Protéine C

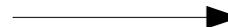


Réaction A
Réaction B
Réaction C

Réaction A :



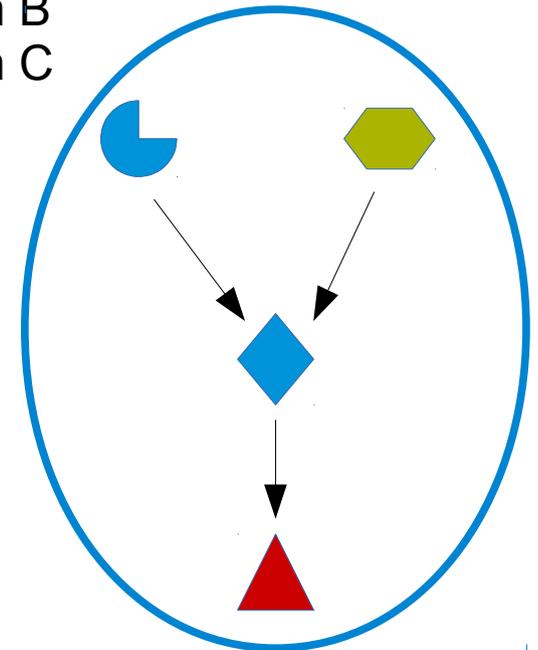
Réaction B :



Réaction C :

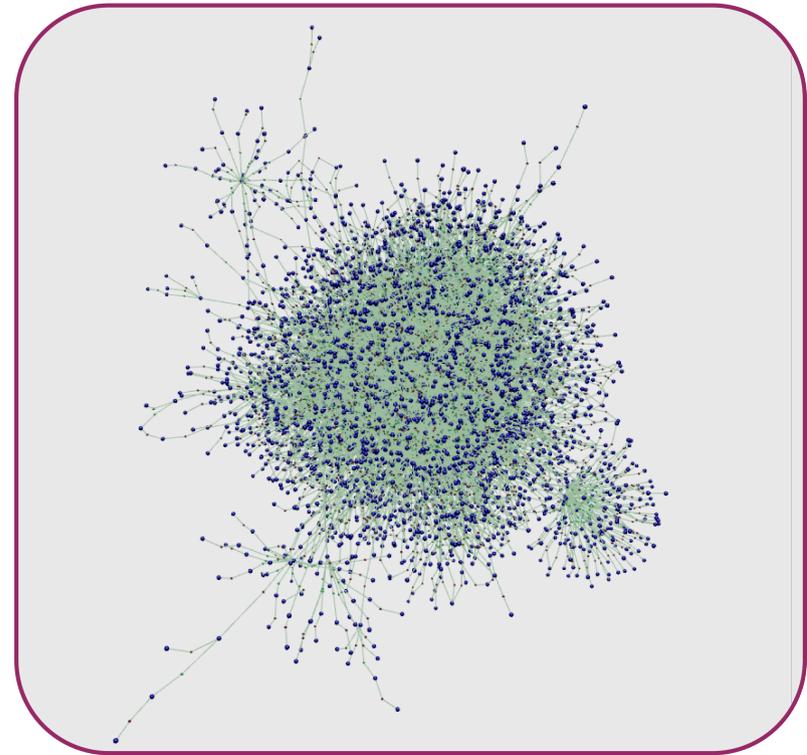


**Réseau
métabolique**



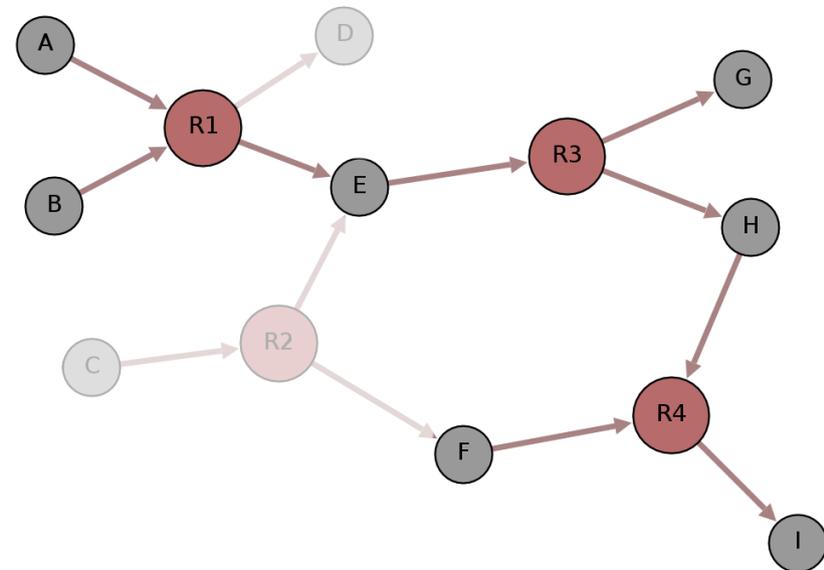
Les réseaux métaboliques

- Évidemment dans la réalité ils sont beaucoup plus gros !!
 - Actuellement pour *E. Siliculosus*
 - 1.844 réactions
 - 2.104 métabolites

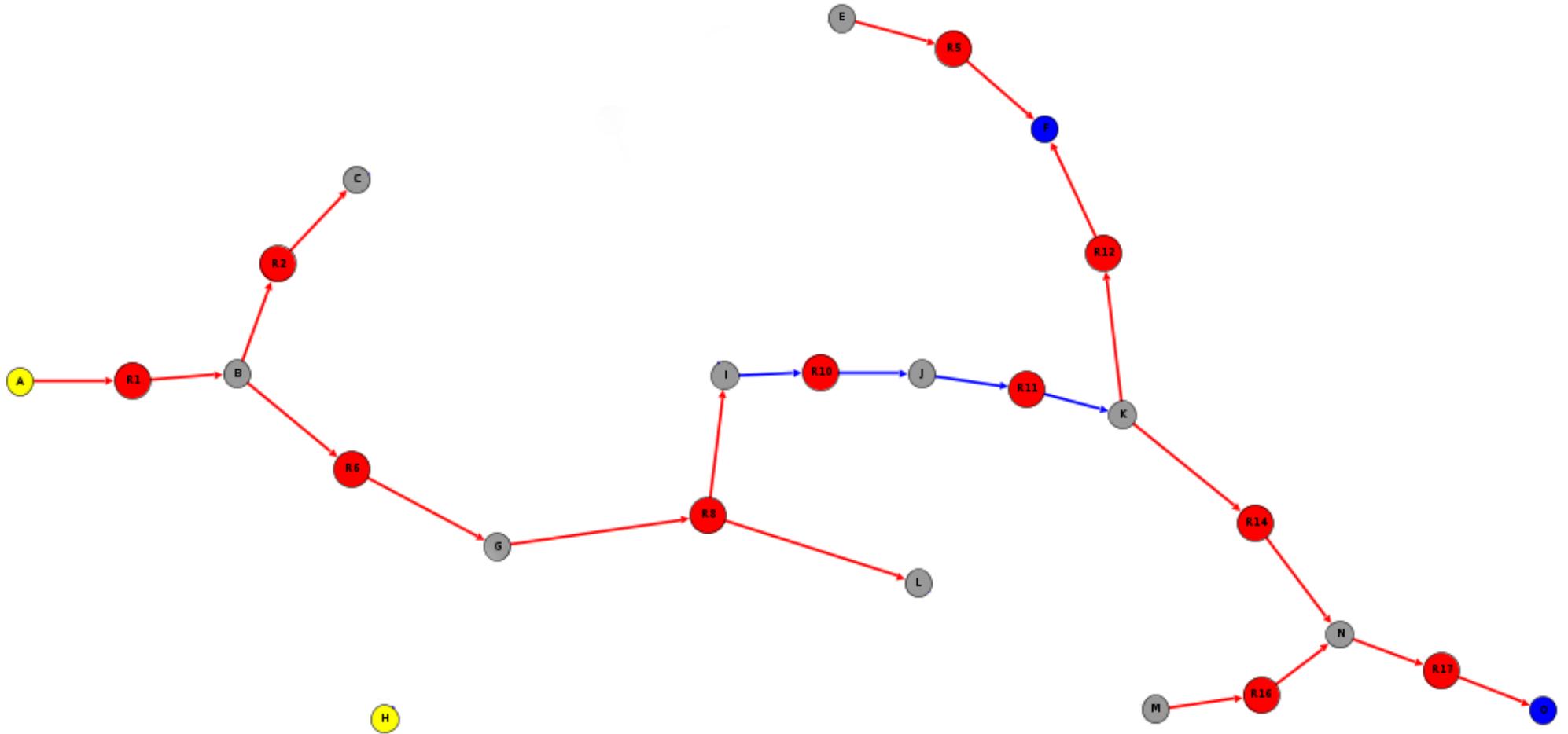


Mon travail réel : reconstruire ce réseau métabolique

- Les informations que l'on a à disposition sont incomplètes, erronées, mal annotées, etc
 - Sur un petit réseau il est facile de trouver ces erreurs à la main
 - Sur un gros réseau, c'est impossible !
- Mais **Meneco** sait le faire !!



Meneco, un logiciel pour boucher des trous



Et maintenant ?

- Étude du réseau obtenu
 - Un mois à Berlin dans un laboratoire spécialisé
- Soutenance de thèse le 30 septembre



Et après?

- Post-doctorat
 - A l'étranger (probablement au Chili)
- Modélisation d'écosystèmes à l'aide de réseaux métaboliques
 - Un organisme ne vit jamais seul !!
 - Prendre en compte l'ensemble des interactions
 - Encore plus de réactions, encore plus composés (ou pas ?), encore plus de fun !

Vous souhaitez en savoir plus sur la bioinformatique ?

- <http://bioinfo-fr.net/>



- Blog bio-informatique

- Réalisé par de jeunes chercheurs en bioinfo
- Blog francophone

- Mail :

- sylvain.s.prigent@irisa.fr

- Questions précises ou générales ?

- **C'est maintenant !**

Merci



Sylvain.s.prigent@irisa.fr

Inria
INVENTEURS DU MONDE NUMÉRIQUE

Lycée Descartes
Rennes
14 février 2014

www.inria.fr