

# Retour d'expérience sur Django, MongoDB : Application Web ThaliaDB

**Journée Intégration  
Visualisation de données**

26 septembre 2018

Delphine Steinbach, UMR GQE-Le Moulon

# ABI : Atelier de Bioinformatique et d'Informatique

<http://moulon.inra.fr/index.php/fr/equipestransversales/atelier-de-bioinformatiqueetinformatique>

UMR Génétique quantitative et Evolution – Le Moulon

Ferme du Moulon

Plateau de Saclay

Gif sur Yvette

Depuis 2016

**Génétique Quantitative et Évolution  
Le Moulon**

recherches ...

Accueil Plan d'accès Mentions légales Annuaire Connexion Intranet

[Présentation](#)

[Équipes](#)

[Équipes transversales](#)

- [Atelier Cartographie Expression Polymorphisme](#)
- [Atelier de Bioinformatique et d'Informatique](#)
- [CATI](#)
- [Publications](#)

[Installation expérimentale](#)

[Plateforme de Protéomique](#)

[Production scientifique](#)

[Ressources tout public](#)

[Contacter l'administrateur](#)

**Atelier de Bioinformatique et Informatique (ABI)**

Les activités d'ABI portent sur :

- l'analyse de la séquence des génomes
- le développement de logiciels et bases de données pour la génétique
- l'informatique système et réseau.

 Harry	 Johann	 Olivier M
 Mélanie	 Adrien	 Olivier
 Delphine	 Pierre	 Benoît
 Yannick	 Laetitia	 Thierry

Septembre 2018

L'atelier de Bioinformatique et Informatique de GQE-Le Moulon se compose de 3 groupes placés sous la responsabilité hiérarchique d'Olivier Martin :

<b>ABI-AnaSeq</b> Resp. : J. Joets  Harry BELCRAM, AI INRA (50 % ACEP) Johann JOETS, IR INRA (50 % DyGAP)	<b>ABI-SOFT</b> Resp. : D. Steinbach  Yannick de OLIVEIRA, IE INRA Pierre MONTALENT, IE INRA Mélanie POLART-DONAT, IE INRA, CDD Delphine STEINBACH, IR INRA	<b>ABI-SYS</b> Resp. : O. Langella  Adrien FALCE, AI INRA, CDD Benoit JOHANNET, AI INRA Olivier LANGELLA, IR INRA (90 % PAPP0)
---	---	---

ABI est rattaché au LabEx BASC, Biodiversité, Agroécosystèmes, Société, Climat.

Copyright © 2018 . Tous droits réservés.  
Joomla! est un logiciel libre sous licence GNU/GPL.

## Environnement logiciel pour la Recherche, dédié **Génétique, Sélection, Multi-espèces**

**F**indable  
**A**ccessible  
**I**nteroperable  
**R**eusable



Maïs  
Blé  
Nigelle  
Levures

*Mission : application aux thématiques de recherche  
de l'UMR, du département BAP. Enjeux INRA2025,  
Climat, OpenScience*

# ABI-SOFT : Tools

<http://moulon.inra.fr/index.php/fr/equipes/abi/336>

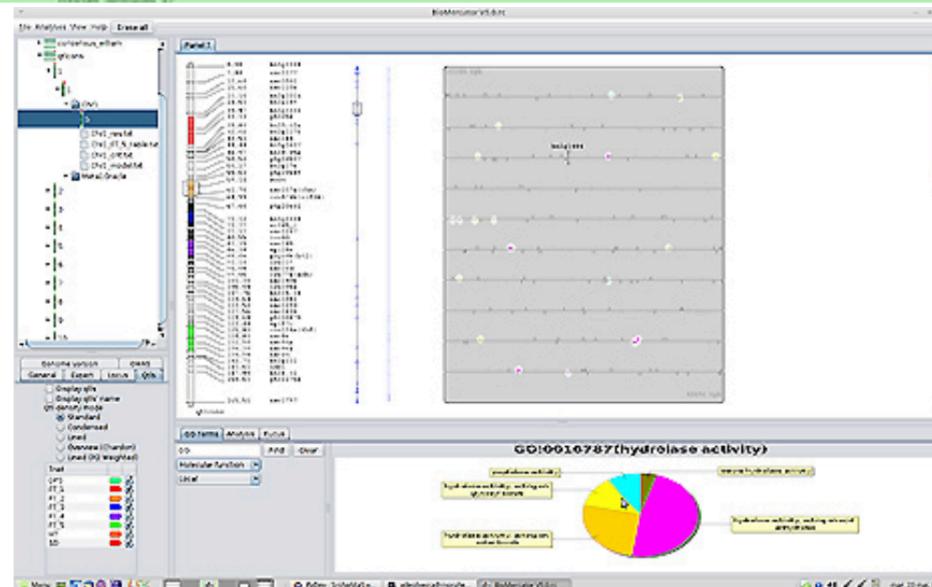


The screenshot shows the Galaxy web interface. On the left, a sidebar lists various tools and workflows. The main content area displays a tool's documentation and its history. A central box contains the text "GALAXY" and "P. Montalent".

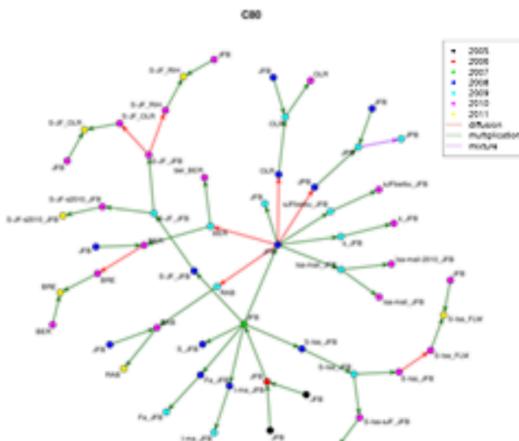
**Biomericator**  
**V5**  
**2018-2019**  

- Méta-QTLs
- GWAS
- Méta-GWAS

**JAVA**  
**Y. De-Oliveira**



**PhenoFieldValidator**  
**2017**  
**Trial : Quality control**  
**Measures, Traits**  
**Rules**  
**Soil, Weather**  
**CSV**  
**ETL : Job Talend TOS**  
**D. Steinbach**



**LandScapeDb : Wheat Diversity data in landscapes**  
 V1 : data France  
 V2: data Europe  
**Collab. DEAP-H2020 RustWatch**

**SHiNeMaS (Seeds History Management and Network System):** relationships between seeds distributed in a network of farms and their field phenotypic evaluation results in several environments (climate, soil)

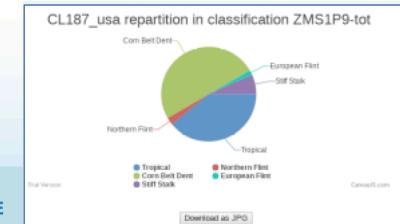
Participative science, agroecology and biological agriculture.  
**Collab. DEAP-RSP**

**Y. De-Oliveira**

## ThaliaDB

A Database for Association Studies

HOME DATAVIEW ADMIN USER PROFILE



ThaliaDB is a database for association studies developed by the [ABI-SOFT team](#) of GQE-Le Moulon in collaboration with the [GQMS team](#), under the project [Amaizing](#).

Diversity and association genetics studies lead to manipulate a large number of individual, lines, clones and/or populations. Moreover, emergence of high-throughput technologies for both genotyping and phenotyping generates a large amount of data. These data need to be stored and managed in order to make requests and to organize datasets to perform genetic diversity data exploration and association genetics analysis. The new version of ThaliaDB, V3.1, is developed for scientists to facilitate their data management and analysis. The resources include access to data, seed lots, samples, genotyping and elaborated phenotyping datasets. It is well adapted for data analysis methods. It can manage high-throughput results coming from different projects and experiments and allows users to explore these data and to give access to them for reuse. This Web tool offers to users a Select (Data selection), Data administration and loading mode. Data confidentiality is maintained using user accounts and specific access rights. It enables data extraction in CSV format. A version exists today in our lab with maize data produced by the GQMS team and their partners. Perspectives are to test this new version for tomato, wheat and poplar. Improvement with funding of Amaizing, Investment for the future, project. It is developed in Python under PostgreSQL and MongoDB databases. Contact: [delphine.steinbach@inra.fr](mailto:delphine.steinbach@inra.fr) for more information and support.



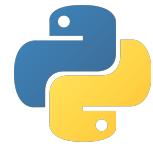
Issoumou, Yannick De Oliveira, <sup>1</sup>Lan-Anh Nguyen, <sup>1</sup>Olivier Akman soy, <sup>1</sup>Delphine Madur, <sup>1</sup>Cyril Baulard, <sup>1</sup>Stéphane Nicolas, <sup>1</sup>Laurence Moreau, <sup>1</sup>Alain Charcosset and <sup>1,2</sup>Delphine Steinbach  
<sup>1</sup>left lab, <sup>2</sup>corresponding author  
 Génétique Quantitative et Evolution – Le Moulon, ferme du Moulin F-91190 Gif-sur-Yvette, France

**ThaliaDB :**  
 V3 : High throughput Genotyping,  
 Phenotyping  
 Genetic resources, SeedLots  
**D. Steinbach**

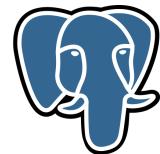
D. Steinbach, ABI-SOFT – 26 sept. 2018

**V3.4**

# ABI-SOFT : Web Tools - Databases



Python



PostgreSQL



Eclipse

Mockflow

1 Serveur SGBD,  
1 Serveur d'applications  
LINUX

**Méthode Agile**

**SourceSup**: Forge

**SharePoint** : ABI-SOFT

**django**

Framework



Contrôleur



Vue

Modèle



Template



# ThaliaDB : DATA

**Gestion de données de plantes et de descripteurs de plantes, d'évaluation de ces plantes dans différents environnements : Genotype X Environnement**

- Lignées, hybrides, collections, populations, panels, de données de structuration de groupes de plantes,
- Échelle locale : parcelle, régionale, pays, continent

Gestion de données de Génotypage **à haut débit** de plantes : multiples expérimentations, multi collaborateurs, multi projets

Gestion de données de mesures de caractéristiques sur des plantes au champ

Gestion de données Images : images de plantes, d'essais

Gestion de confidentialité des données, partiel, totale, gestion des accès : sécurité

# ThaliaDB : Besoins en 2015

## Existant : ThaliaDB V2 (JAVA)

- Design graphique obsolète
- Requêtes à éviter
- *Bugs et pas que mineurs..., mais connus*
- Ne permettait pas l'ajout de nouveaux jeux de données nouvelles technologies haut débit : génotypages (maïs)
- *Compliquée à maintenir, à faire évoluer sur le long terme, manque de RH spécialisé en JAVA.*
- *Fonctionnement avec stagiaires, CDD*

## Remise à plat

- Du projet
- De l'organisation : RH, Hardware, Software
- Etude des besoins d'une V3 : feuille de route
  - Nouveaux développements
  - Nouvelles insertions de données par les biologistes

# ThaliaDB : Besoins

Accessible, sécurisé depuis l'extérieur : <https://thalia.moulon.inra.fr>

- Gestion fine des accès, des comptes utilisateurs, de droits d'accès à différents jeux de données (hiérarchie)

Administration des données par les biologistes en autonomie

Developpement : Evolutivité, Transférable

- à d'autres équipes
- unités du département BAP, EFPA
- tiers, collaborateurs
- Autres espèces de plantes (ex: blé, tomate, pois, colza, pommier, rosier, peuplier, chêne)

Documentation

- Guide d'installation, Guide d'utilisation
- Forge : SourceSup

# ThaliaDB et le BigData

## Intégration, visualisation, exploration : multiples jeux de données, multi projets

- Diverses sources
- Structurés, moins structurés
- Différents délais d'arrivées
- Différents protocoles d'analyse en amont de la base et en aval

## Requêtes sur des objets d'intérêt scientifiques

- Formulaires **d'administration des données** : ins/upd/del
- Formulaires de **requêtes multi-critères** (Select)
- **Outils de filtrage** : différents niveaux **d'aggrégation**
- **Outils d'export de données** : différents formats de sortie : standards, pour des entrées de logiciels (workflow)
- **Outils graphiques (dynamique) de visualisation des données**

# ThaliaDB : Besoins

## Partage, Export, Inter-opérabilité

- Entrepots de PF (GnplS)
- Nationaux: Data.inra.fr, européen
- International : MaizeGDB, ...

## Types de données

- Brutes : \* priorité \*
- Elaborées:
  - analysées (calculs, moyennes, imputations...) : différentes méthodes

## Cas d'utilisation : maïs

- 20 ans de données du laboratoire : équipe GQMS, A. Charcosset
- ses projets en partenariat
  - Ex : PIA Amaizing
  - Ex: projet européen DROPS
  - Cibles : sécheresse, froid, Azote

## Exemple : jeux de données

- Génotypage de la puce 50 000 marqueurs
  - 17 expériences => 17 matrices
- Génotypage de la puce : 600 000 marqueurs : 3 expériences
- Génotypage par séquençage : 900 000 marqueurs
- Nouvelles données de phénotypage au champs
  - Réseau d'essais au champ : 8 sites, 5 panels de plus de 300 plantes de n projets

## Croisement de jeux de données de génotypage et de phénotypage

- Valeur ajoutée

## Apporter des données pour prédire : valeur ajoutée, nouvelles données

- Construire des expérimentations de sélection génomique
  - Prédictions de phénotypage à partir de génotypage et de phénotypage
- Fournir des données à des modèles d'éco-physiologie
  - Prédiction comportements de plantes dans différents environnements
- Fournir des données pour des modèles statistiques: deep learning, ...

# ThaliaDB : Solution

## Architecture

- **Python : 3.5.3, panda**
- **Framework Django : harmonisation (applications de l'équipe) : 1.10** (passage en 2.0 prévu) : jsonfield, mongoengine, autocomplete
- **Apache : 2.2**
- **SGBDR : PostGreSQL : 9.6.7**
  - Utilisation de PG JSON pour certains champs
- **NoSQL : veille**
  - **MongoDb : BD orientée Documents : 3.2.19, free, openSource,** <https://www.mongodb.com/what-is-mongodb>
  - **Analyse comparative entre ElasticSearch (moteur de recherche) et MongoDB** faite dans CATI CGI entre GnpIS et BD Cirad : données de génotypage
    - Résultats quasi-similaires, dépendantes des requêtes et des besoins des applicatifs Web
    - Participation aux MongoDB Days, Paris
  - **Décision : tester mongoDB pour ThaliaDB**
    - Hypothèse *maîtrisable* dans notre contexte local, faible RH, % ETP non permanents élevé
  - **ThaliaDB : BD Hybride : PostGreSQL et MongoDB**

# Framework Django :

## <https://www.djangoproject.com/>



Django makes it easier to build better Web apps more quickly and with less code.

[Get started with Django](#)

## Meet Django

Django is a high-level Python Web framework that encourages rapid development and clean, pragmatic design. Built by experienced developers, it takes care of much of the hassle of Web development, so you can focus on writing your app without needing to reinvent the wheel. It's free and open

D. Steinbach, ABI-SOFT – 26 sept. 2018

[Download latest release: 2.1.1](#)

[DJANGO DOCUMENTATION ›](#)

Support Django!

# Framework Django :

## Tutoriel + Tutoriel sur OpenClassrooms

← → ⌂ 🔒 https://docs.djangoproject.com/en/2.1/ ⭐ 🙏 🔍

### First steps

Are you new to Django or to programming? This is the place to start!

- **From scratch:** [Overview](#) | [Installation](#)
- **Tutorial:** [Part 1: Requests and responses](#) | [Part 2: Models and the admin site](#) | [Part 3: Views and templates](#) | [Part 4: Forms and generic views](#) | [Part 5: Testing](#) | [Part 6: Static files](#) | [Part 7: Customizing the admin site](#)
- **Advanced Tutorials:** [How to write reusable apps](#) | [Writing your first patch for Django](#)

### The model layer

Django provides an abstraction layer (the “models”) for structuring and manipulating the data of your Web application. Learn more about it below:

- **Models:** [Introduction to models](#) | [Field types](#) | [Indexes](#) | [Meta options](#) | [Model class](#)
- **QuerySets:** [Making queries](#) | [QuerySet method reference](#) | [Lookup expressions](#)
- **Model instances:** [Instance methods](#) | [Accessing related objects](#)
- **Migrations:** [Introduction to Migrations](#) | [Operations reference](#) | [SchemaEditor](#) | [Writing migrations](#)
- **Advanced:** [Managers](#) | [Raw SQL](#) | [Transactions](#) | [Aggregation](#) | [Search](#) | [Custom fields](#) | [Multiple databases](#) | [Custom lookups](#) | [Query Expressions](#) | [Conditional Expressions](#) | [Database Functions](#)
- **Other:** [Supported databases](#) | [Legacy databases](#) | [Providing initial data](#) | [Optimize database access](#) | [PostgreSQL specific features](#)

### The view layer

You are here:

- [Django 2.1 documentation](#)
  - Django documentation

### Getting help

#### FAQ

Try the FAQ – it’s got answers to many common questions.

[Index](#), [Module Index](#), or [Table of Contents](#)  
Handy when looking for specific information.

#### [django-users mailing list](#)

Search for information in the archives of the django-users mailing list, or post a question.

#### #django IRC channel

Ask a question in the #django IRC channel, or search the IRC logs to see if it’s been asked before.

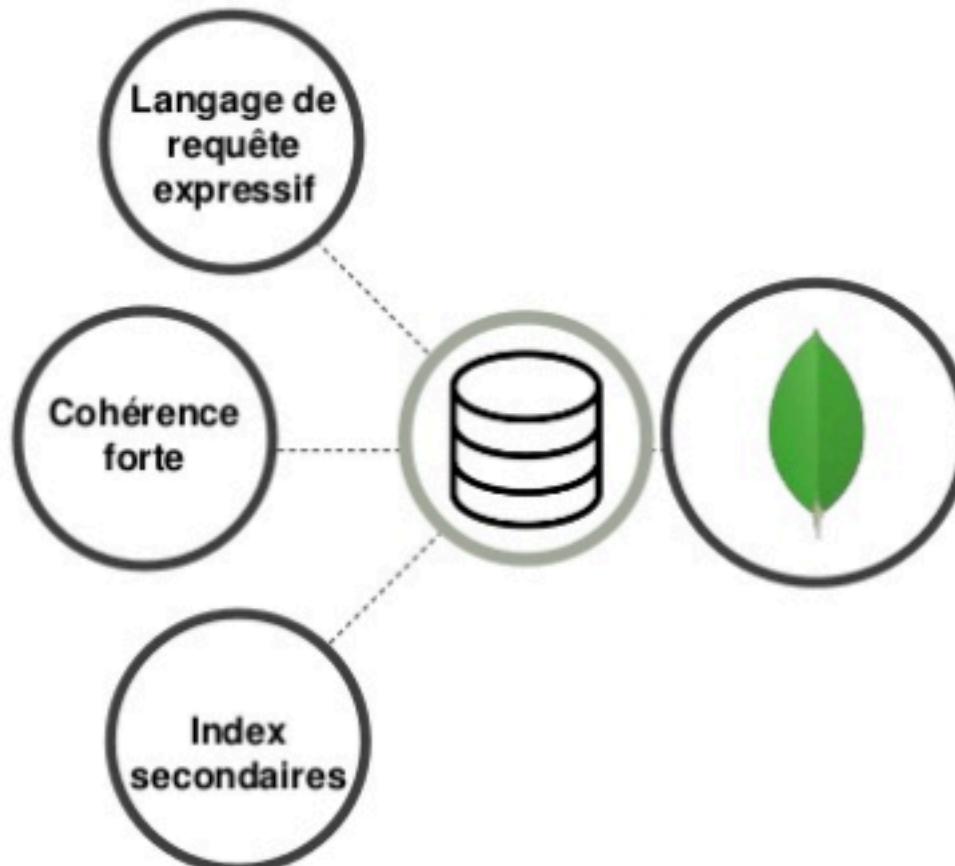
Language: [en](#)

#### [Ticket tracker](#)

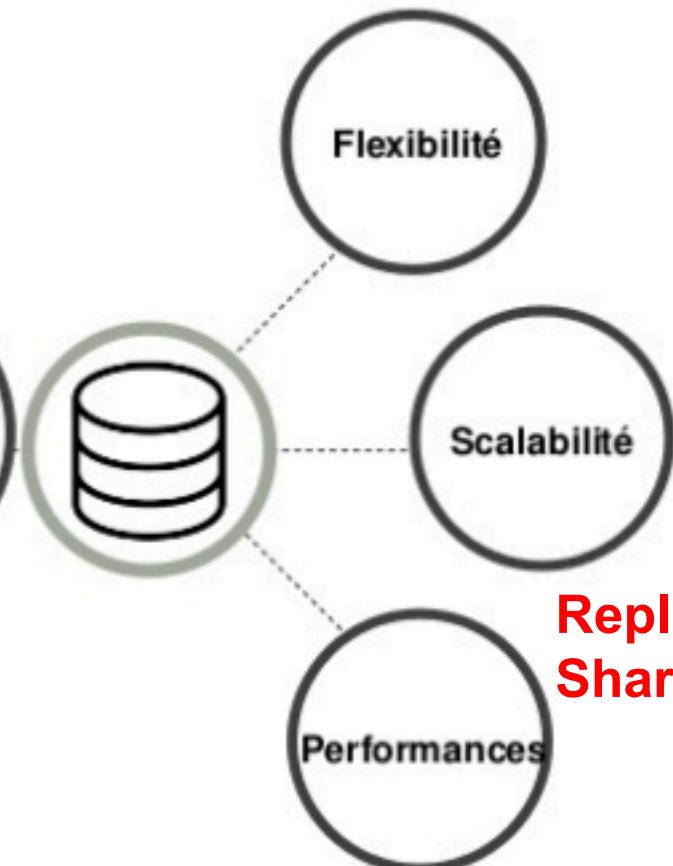
Report bugs with Django or documentation version: 2.1  
in our ticket tracker.

# NoSQL - SGBDR

## Base de données relationnelle



## NoSQL



Base de données relationnelle + NoSQL

 mongoDB

# MONGODB :

<https://www.mongodb.com/what-is-mongodb> :  
clients

← → C https://www.mongodb.com/what-is-mongodb ☆ ⌂

DOCS LEARN WHAT'S MONGODB? BLOG LOGIN

Get MongoDB ⬇

mongoDB. | FOR GIANT IDEAS Trusted by thousands of teams SOLUTIONS CLOUD CUSTOMERS RESOURCES ABOUT US

Adobe AMADEUS amazon.com BOSCH AstraZeneca

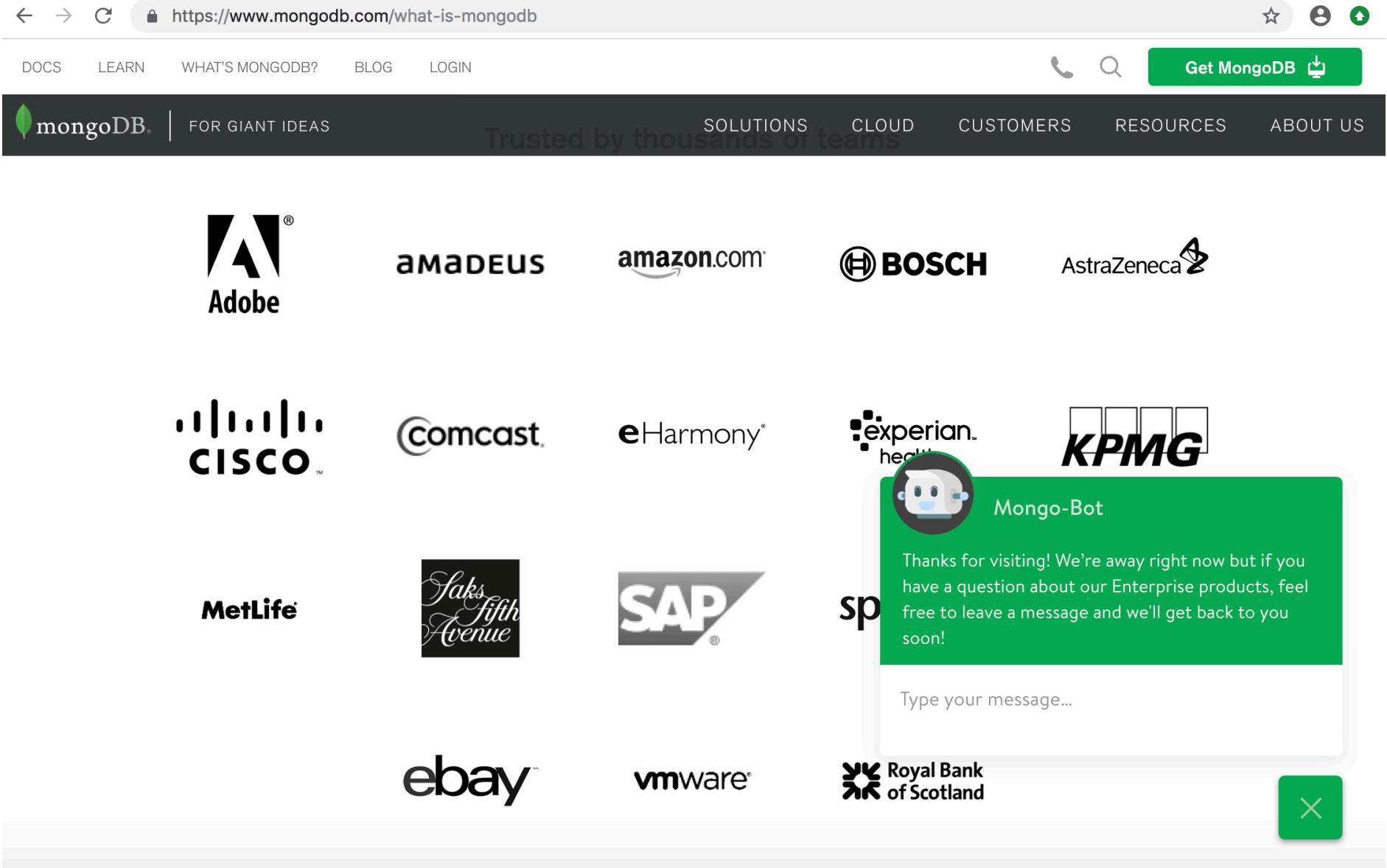
CISCO Comcast eHarmony experian health KPMG

MetLife Saks Fifth Avenue SAP sp Mongo-Bot

Thanks for visiting! We're away right now but if you have a question about our Enterprise products, feel free to leave a message and we'll get back to you soon!

Type your message...

ebay vmware Royal Bank of Scotland X



# MONGODB : Documentations

← → ⌂ https://docs.mongodb.com/?\_ga=2.92647956.845964873.1538025228-1825171843.1538025228

 mongoDB | Documentation ▾ Search Documentation

**Overview:**

-  Getting Started
-  Guides
-  MongoDB Products
  -  MongoDB Server
  -  MongoDB Stitch
  -  MongoDB Drivers
-  Cloud
  -  MongoDB Atlas
  -  MongoDB Cloud Manager
  -  MongoDB Ops Manager
-  Tools
  -  MongoDB BI Connector
  -  MongoDB Charts
  -  MongoDB Compass
  -  MongoDB Spark Connector

Questions? Contact Us 

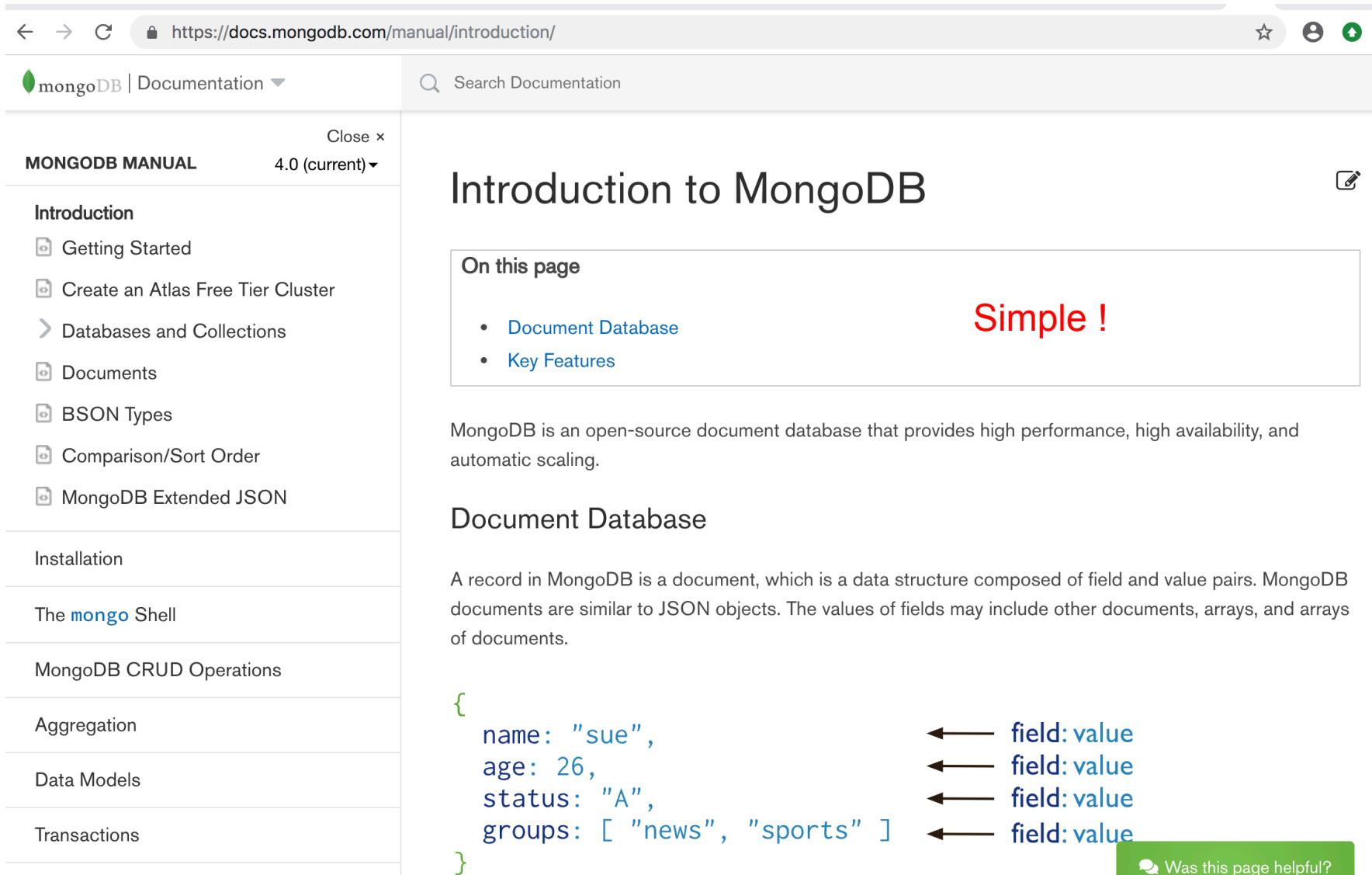
## Welcome to the MongoDB Docs

 **Getting Started**  
Connect and use MongoDB with your favorite programming language within minutes.

 **MongoDB Manual**  
The MongoDB Manual, including: CRUD, aggregation, database commands, mongo shell methods, and other common operations in MongoDB.

 **Cloud Products**  
MongoDB offers a variety of cloud

# MONGODB : Database = Collection



The screenshot shows a browser window displaying the MongoDB Documentation website at <https://docs.mongodb.com/manual/introduction/>. The page title is "Introduction to MongoDB". On the left, there is a navigation sidebar with the following sections:

- MONGODB MANUAL** (Version 4.0 (current))
- [Introduction](#)
- [Getting Started](#)
- [Create an Atlas Free Tier Cluster](#)
- [Databases and Collections](#)
- [Documents](#)
- [BSON Types](#)
- [Comparison/Sort Order](#)
- [MongoDB Extended JSON](#)
- [Installation](#)
- [The `mongo` Shell](#)
- [MongoDB CRUD Operations](#)
- [Aggregation](#)
- [Data Models](#)
- [Transactions](#)

The main content area features the heading "Introduction to MongoDB" with a red "Simple !" callout. Below it is a section titled "On this page" with two links: "Document Database" and "Key Features". A large text block describes MongoDB as an open-source document database. To the right, a code snippet illustrates a MongoDB document structure:

```
{  
  name: "sue",  
  age: 26,  
  status: "A",  
  groups: [ "news", "sports" ]  
}
```

Annotations with arrows point from the text "field: value" to each field-value pair in the document: "name: "sue",", "age: 26,", "status: "A",", and "groups: [ "news", "sports" ]".

**Was this page helpful?**

# MONGODB : Insert Documents : JSON

← → ⌂ https://docs.mongodb.com/manual/tutorial/getting-started/ ⭐ 🙏 🔍

 MongoDB | Documentation ▾

MONGODB MANUAL Close × 4.0 (current) ▾

- Introduction
- Getting Started**
- Create an Atlas Free Tier Cluster
- Databases and Collections
- Documents
- BSON Types
- Comparison/Sort Order
- MongoDB Extended JSON

---

- Installation
- The `mongo` Shell
- MongoDB CRUD Operations
- Aggregation
- Data Models
- Transactions
- Indexes

Search Documentation

- **Local MongoDB installation.** For more information on installing MongoDB locally, see [Install MongoDB](#).

## Insert Documents

`pymongo.collection.Collection.insert_many()` can insert *multiple documents* into a collection. Pass an array of documents to the method.

The following example inserts new documents into the `inventory` collection:

```
db.inventory.insert_many([
    # MongoDB adds the _id field with an ObjectId if _id is not present
    { "item": "journal", "qty": 25, "status": "A",
      "size": { "h": 14, "w": 21, "uom": "cm" }, "tags": [ "blank", "red" ] },
    { "item": "notebook", "qty": 50, "status": "A",
      "size": { "h": 8.5, "w": 11, "uom": "in" }, "tags": [ "red", "blank" ] },
    { "item": "paper", "qty": 100, "status": "D",
      "size": { "h": 8.5, "w": 11, "uom": "in" }, "tags": [ "red", "blank", "plain" ] },
    { "item": "planner", "qty": 75, "status": "D",
      "size": { "h": 22.85, "w": 30, "uom": "cm" }, "tags": [ "blank", "red" ] },
    { "item": "postcard", "qty": 45, "status": "A",
      "size": { "h": 10, "w": 15.25, "uom": "cm" }, "tags": [ "blue" ] }
])
```

**copy**

`insert_many()` returns a document that includes the newly inserted document.

Was this page helpful? 

# ThaliaDB et MongoDB

## Résultats

- Création d'un schéma pour la gestion du génotypage compatible avec les autres données sous MongoDB :
    - +- imbrication
  - Test de requêtes en connection directe sur MongoDB
    - Jusque 5 millions de marqueurs
  - Test de requêtes via framework Django
  - Serveur de développement
  - Serveur de production
  - PC local
- 
- Achat d'un disque dédié sur le serveur pour MongoDB : 1 To
  - Test d'insertion de données à haut débit : 50k, 600K, 1Million
  - Outils d'admin, monitoring (gratuits)

## Résultats

- Conservation de PostGreSQL pour garder cohérence sur objets phares
- Ajout de MongoDB :
  - pour gestion du génotypage à haut débit
  - = modèle HYBRIDE !
  - Outil de supervision : Munin
- Passage en production depuis automne 2017
  - Avec migration des données de la V2 !
  - Ouverture de comptes pour les administrateurs de données (biologistes)
  - Nouvelles données haut débit arrivées ! : 15 Go génotypage
    - **35 expériences**
      - sur puce 50 000 marqueurs
      - sur puce 600 000
    - GBS : 1 million de marqueurs : *en cours en production*
  - **6000 plantes, 30 projets, 80 comptes utilisateurs, 32 environnements de phénotypage, 104 traits mesurés**

# ThaliaDB : Modèle Hybride

Les +

- **V3 marche !**
- **Bugs résiduels ‘connus’ en cours de correction, Ajout de tests dans le code**
- **Evolutif** : ajout de jeux de données de petite et de grande taille
- **Flexible, Agile**
- **Haute disponibilité**
- **Volumétrie**
- **Schéma dynamique : imbrication de niveaux !**
- **Performances élevées** : encore plus hors framework
- **Apprentissage assez facile**, pas trop de difficulté par rapport à un modèle relationnel
- **Outils: Mongoimport, Mongodump**
- **Potentialité de requêtes** à l'avenir encore plus enrichies, geospatial, de texte, agrégation,
- **Mode Hybride : Relationnel – NoSQL** : combine les avantages
- **Pas encore besoin de réPLICATION**, pas de demande *payante* à MongoDB à ce stade
- **Formation MongoDB** : utile pour aggréger documentation et me remettre dans le bain ‘admin’
- **MongoDb Days** : utile : retours d’expériences des privés et des gros utilisateurs
- **Communauté Mongo** : Conseils, offre de formation, version Enterprise !

Les -

- **La mémoire a saturé une fois !** (mémoire était partagée avec d'autres appli)
  - Lié à applicatif ? à Mongo ?
- **Outil de supervision mis en place** : Graphique : Accès mémoire, disques, réseaux
- **Amélioration du code des requêtes** : optimisation, pb non reproduit à ce jour mais pas de perte de données.
  - **Préconisation : serveur dédié**
- **Un time out coté Web** pour l'insertion via l'appli de la méga matrice....
  - Changement d'un paramètre sur Django
- **Moins rapide via framework** (= effet framework)
- **Cout de la formation MongoDB mais valait le coup pour synthétiser tout en peu de jours..., et prévoir en cas de besoin de scalabilité et d'architecture plus complexe**
- **Offre Mongo** : Conseils, offre de formation (payante), version Enterprise (payante) pour avoir les outils adaptés à **une production de très haute dispo.**
  - Scalabilité ne serait pas 100% mais pas testé.
- **Le 'En cours' : étude de la gestion des mises à jour des documents MongoDB, Optimisation**

**ANNEXES** : extraits d'une présentation : <https://www.slideshare.net/mongodb/quand-utiliser-mongodb-et-quand-vous-en-passer>

The image shows a presentation slide with a dark background. At the top, the title 'Quand utiliser MongoDB' is displayed in large white font. Below it, the subtitle '... Et quand vous en passer...' is also in white. In the center, there is author information: 'Docteur Edouard Servan-Schreiber', 'Directeur d'architecture de solution', 'MongoDB', and the email 'edss@mongodb.com'. At the bottom of the slide, there is a navigation bar with a blue progress bar, several small white dots, and standard presentation controls (back, forward, search).

Quand utiliser MongoDB  
... Et quand vous en passer...

Docteur Edouard Servan-Schreiber  
Directeur d'architecture de solution  
MongoDB  
edss@mongodb.com

◀ 2 of 46 ▶ 🔍

# ThaliaDB

A Database for Association Studies

HOME    DATAVIEW    ADMIN    USER PROFILE    HELP    LOGOUT



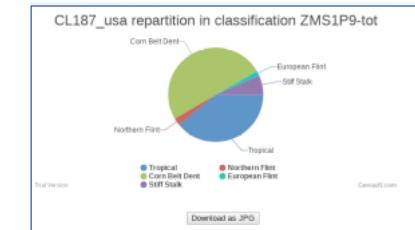
ThaliaDB is a database for association studies developed by the [ABI-SOFT team](#) of [GQE-Le Moulon](#) in collaboration with the [GQMS team](#), under the project [Amaizing](#).

Diversity and association genetics studies lead to manipulate a large number of individual, lines, clones and/or populations. Moreover, emergence of high-throughput technologies for both genotyping and phenotyping generates a large amount of data. These data need to be stored and managed in order to make requests and to organize datasets to perform genetic diversity data exploration and association genetics analysis. The new version of ThaliaDB, V3.1, is developed for scientists to facilitate their data management and analysis. The database holds genetic resources data, seed lots, samples, genotyping and elaborated phenotyping datasets. It is well adapted for data that are useful to apply GWAS methods. It can manage high-throughput results coming from different projects and experiments and propose several views and options to explore these data and to give access to them for reuse. This Web tool offers to users a Select (Data view) mode and an Admin (Data administration and loading) mode. Data confidentiality is maintained using user accounts and specific levels of rights can be set on data. It enables data extraction in CSV format. A version exists today in our lab with maize data produced from projects of A. Charcosset's GQMS team and theirs partners. Perspectives are to test this new version for tomato, wheat and poplar. The software is currently in improvement with funding of Amaizing, Investment for the future, project. It is developed in Python under Framework Django, running under PostGreSQL and MongoDb databases. Contact: [delphine.steinbach@inra.fr](mailto:delphine.steinbach@inra.fr) for more information and collaboration.

<sup>1</sup>Alice Beaugrand, <sup>1</sup>Guy-Ross Assoumou, <sup>1</sup>Yannick De Oliveira, <sup>1</sup>Lan-Anh Nguyen, <sup>1</sup>Olivier Akmansoy, <sup>1</sup>Delphine Madur, <sup>1</sup>Cyril Bauland, <sup>1</sup>Stéphane Nicolas, <sup>1</sup>Laurence Moreau, <sup>1</sup>Alain Charcosset and <sup>1,2</sup>Delphine Steinbach

<sup>1</sup>left lab, <sup>2</sup>corresponding author

<sup>1</sup>Génétique Quantitative et Evolution – Le Moulon, ferme du Moulon F-91190 Gif-sur-Yvette, France



Dépôt projet CATI 3G :  
**PlantBreed**  
20 bioinformaticiens, BAP

## Merci pour votre attention Questions ?