Les bases de données en bioinformatique

Jean-Baptiste Lamy

Maître de conférence
bureau 149, LIM&BIO
jibalamy@free.fr

Introduction

- Enseignement :
  - 3 x 2h de cours
  - 3 x 4h de TP en 3 groupes
- Contenu :
  - Introduction
  - Créer et remplir une base de données
  - Faire des recherches dans une base de données (requêtes)
  - Les grandes bases de données existantes en bioinformatique

Définitions

- **Donnée :**
  - Donnée = une valeur (avec ses unités, son nom,...)
  - Ex : la glycémie de monsieur X est de 1,5g/l
- **Information :**
  - Information = donnée + sens
  - C'est une donnée associée à son interprétation
  - Ex : la glycémie de monsieur X est de 1,5g/l, donc monsieur X est diabétique
- **Connaissance :**
  - Connaissance = information suffisamment générale pour être réutilisée dans un autre contexte
  - Ex : si la glycémie est supérieure à 1g/l, le patient est diabétique
    - => s'applique à monsieur X mais aussi aux autres patients
- **Informatique :**
  - Informatique = science du traitement **automatique** des données, afin de générer de l'information et de la connaissance

Définitions

- **Donnée, information ou connaissance ?**
  - Il fait froid dehors, le thermomètre indique -2°C !
  - Ces raviolis doivent être cuits pendant 8 minutes
  - L'ARN polymérase se fixe au niveau de la boîte TATA

Définitions

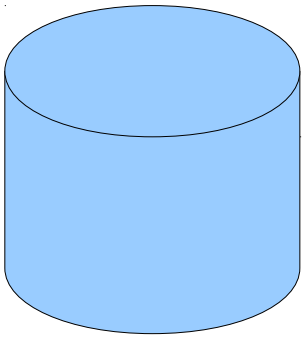
- **Donnée, information ou connaissance ?**
  - Il fait froid dehors, le thermomètre indique -2°C !
    => Information
  - Ces raviolis doivent être cuits pendant 8 minutes
    => Donnée
  - L'ARN polymérase se fixe au niveau de la boîte TATA
    => Connaissance

Systèmes d'information

- **Système d'information :**
  - Système permettant de regrouper, classer et diffuser de l'information
    - **Objectif : donner la bonne information à la bonne personne et au bon moment**
  - Exemples :
    - Gestion humaine
    - Gestion documentaire
    - Gestion d'un stock de marchandise
    - Gestion des remboursements par la sécurité sociale
    - ...
  - Aujourd'hui de plus en plus des systèmes informatiques
  - Mais pas nécessairement
    - Ex : carnet de santé

Base de données

- Base de données : données organisées et structurées de sorte à pouvoir stocker et manipuler de **grandes quantités** de données
- Les données sont organisées selon un **modèle de données**
  - Ex : dans une base de données de gènes, chaque gène est structuré de la même manière, avec par exemple son nom, sa séquence, etc
- Abréviation BD (*database*, DB en anglais)
- Sur les schémas, les bases de données sont représentés par des cylindres



Base de données

- Un **système de gestion de base de données** (SGBD) est un logiciel qui gère les données suivant un modèle précis, et s'occupe de les traduire en fichiers informatiques de manière optimale
  - Gestion du réseau : bases de données accessibles via Internet,...
  - Gestion de plusieurs utilisateurs simultanément
    - Que se passe-t-il si deux utilisateurs veulent modifier la base de données au même moment ?
  - Gestion des **transactions**
    - Que se passe-t-il s'il y a une panne de courant juste au moment où un utilisateur modifiait la base de données ?
- Il existe de nombreux logiciels de ce type
- Attention ! Des logiciels comme Excel **ne sont pas** des bases de données mais des tableurs !
  - Nombre de lignes limité à 65 000,...
  - Mais plus de 150 000 000 de séquences biologiques (ADN, protéines,...) connues aujourd'hui !

Base de données relationnelle

- Il existe plusieurs catégories de modèles de données
- Le plus fréquent est celui des base de données **relationnelles**
  - Il s'appuie sur l'algèbre relationnelle
- Un modèle relationnel comprend :
  - Des tables
    - Ex : une table de gènes, une table de protéines, une table d'espèces,...
  - Des relations entre ces tables
    - Ex : la table « gène » est reliée à la table « protéine » par une relation qui indique quel gène code quelle protéine

Base de données relationnelle

- **Une table :**
  - Un tableau avec des lignes et des colonnes
  - Chaque ligne correspond à un individu / un élément (un patient, un gène, une protéine,... on parle d'**entités**)
    - Une table ne contient qu'un seul type d'individu / d'élément : que des patients, que des gènes,...
      - => on nomme la table en fonction de ce que contiennent ses lignes : table des patients, table des gènes,...
  - Chaque colonne correspond à un **attribut** (ou propriété, caractéristique, champ,...) : l'âge des patient, la séquence d'un gène,...
    - Un attribut est associé à un type de données :
      - L'âge est un nombre entier
      - La séquence est du texte (« **chaîne de caractère** »)
      - ...

Base de données relationnelle

- Exemple de table :

Table Patient

Identifiant	Nom	Date de naissance	Taille	Poids
1	LAMY	18/06/1979	170	62
2	X	25/03/1957	165	54
3	Y	04/01/1982	180	90

Base de données relationnelle

- Exemple de table :

Table

Identifiant	Nom	Longueur	Séquence
1	Aldostérone synthase	1512	atggcactcagggcaaaggcaga...
2	Enveloppe du VIH	105	tgtacaagacccaacaacaacacaa...
3	Insuline	990	MALWMRLLPLLALLALWG...

Base de données relationnelle

- Exemple de table :

Table

Identifiant	Nom	Longueur	Séquence
1	Aldostérone synthase	1512	atggcactcagggcaaaggcaga...
2	Enveloppe du VIH	105	tgtacaagacccaacaacaacacaa...
3	Insuline	990	MALWMRLLPLLALLALWG...

Base de données relationnelle

- Exemple de tables :

Table gène

Identifiant	Nom	Longueur	Séquence
1	Aldostérone synthase	1512	atggcactcagggcaaaggcaga...
2	Enveloppe du VIH	105	tgtacaagacccaacaacaacacaa...

Table protéine

Identifiant	Nom	Longueur	Séquence
1	Insuline	990	MALWMRLLPLLALLALWG...

Base de données relationnelle

- **Une relation :**
  - Relie deux tables entre elles
  - **Relation 1-1** :
    - on relie chaque élément de la première table à un seul élément de la seconde
    - on relie chaque élément de la seconde table à un seul élément de la première

Table patient

Identifiant	Nom	Date de naissance
1	LAMY	18/06/1979
2	X	25/03/1957
3	Y	04/01/1982

1 A pour lit 1

Table lit

ID	Etage	Service
1	1	Cardiologie
2	1	Pneumologie

Base de données relationnelle

- **Une relation :**
  - Relie deux tables entre elles
  - **Relation 1-*** :
    - on relie chaque élément de la première table à zéro, un ou plusieurs éléments de la seconde
    - on relie chaque élément de la seconde table à un seul élément de la première

Table espèce

ID	Nom	Nom latin
1	Homme	Homo sapiens
2	Poisson zèbre	Danio rerio

1 A pour gène ***

Table gène

ID	Nom	Séquence
1	Aldostérone synthase	atggcactca...
2	Enveloppe du VIH	tgtacaagaccc...

Base de données relationnelle

- **Une relation :**
  - Relie deux tables entre elles
  - **Relation *-1** :
    - on relie chaque élément de la première table à un seul élément de la seconde
    - on relie chaque élément de la seconde table à un ou plusieurs éléments de la première

Table codon

ID	Séquence
1	AUG
2	AUC
3	AUU

* Code 1

Table acidé aminé

ID	Nom	Code
1	Cystéine	C
2	Alanine	A

Base de données relationnelle

- **Une relation :**
  - Relie deux tables entre elles
  - **Relation *-*** : on relie un ou plusieurs éléments de la première table à un ou plusieurs éléments de la seconde

Table espèce

ID	Nom	Nom latin
1	Homme	Homo sapiens
2	Poisson zèbre	Danio rerio

* Vit dans *

Table environnement

ID	Nom	Climat
1	Forêt tropicale	Tropical
2	Jungle	Equatorial

Base de données relationnelle

- Il est possible de relier une table avec elle-même !

Table personne

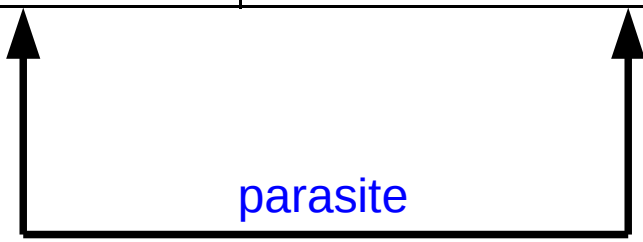
Identifiant	Nom	Date de naissance	Taille	Poids
1	LAMY	18/06/1979	170	62
2	X	25/03/1957	165	54
3	Y	04/01/1982	180	90

Base de données relationnelle

- Il est possible de relier une table avec elle-même !
  - De quelle type de relation s'agit-il ?

Table espèce

ID	Nom	Nom latin
1	Homme	Homo sapiens
2	Palludisme	Plasmodium falciparum

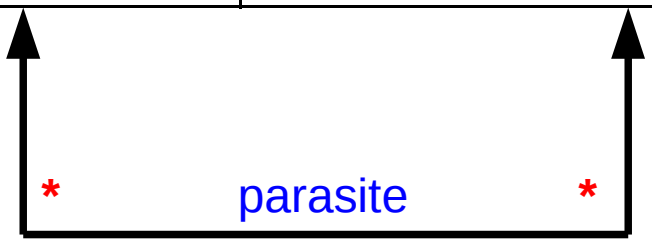


Base de données relationnelle

- Il est possible de relier une table avec elle-même !

Table espèce

ID	Nom	Nom latin
1	Homme	Homo sapiens
2	Palludisme	Plasmodium falciparum



Base de données relationnelle

- Un **système de gestion de base de données relationnelles** (SGBDR) est SGBD qui gère des bases de données de type relationnel
- Quasiment tous les SGBDR partagent le même langage : **SQL (Structured Query Language)**
  - Créé en 1970 par E.F. Codd puis normalisé en 1986
  - Version 3 : 1999
- Les deux principaux SGBDR utilisés « en vrai » :
  - Oracle
    - **TRES** cher !
    - A réserver aux bases de données **TRES** volumineuse
  - **MySQL**
    - Logiciel libre, gratuit
    - Très utilisé sur Internet
    - => On l'utilisera en TP

MySQL

- Logiciel en ligne de commande
  - On tape des commandes qui sont des ordres donnés à l'ordinateur
  - Appuyer sur entrée pour valider les commandes
  - Appuyer sur les flèches haut et bas pour remonter l'historique des commandes et retrouver les commandes tapées précédemment
- Installation sous windows :
  - Inclut dans EasyPHP
    - http://www.easyphp.org/

Créer et remplir une base de données

Comment construire une base de données en biologie ?

- Deux étapes :
  - **1) définir le modèle de données :**
    - Définir les tables et les relations
    - Les décrire en SQL
    - Cette étape est délicate !
    - Si le modèle n'est pas bon, on sera bloqué à l'étape 2 !

Table espèce

ID	Nom	Nom latin

1 A pour gène *

Table gène

ID	Nom	Séquence

Comment construire une base de données en biologie ?

- Deux étapes :
  - **2) remplir la base de données selon le modèle défini :**
    - Cette étape peut être très longue selon la taille de la base de données !

Table espèce

ID	Nom	Nom latin
1	Homme	Homo sapiens
2	Poisson zèbre	Danio rerio

1 A pour gène *****

Table gène

ID	Nom	Séquence
1	Aldostérone synthase	atggcactca...
2	Enveloppe du VIH	tgtacaagaccc...

Comment construire une base de données en biologie ?

- Comment définir le modèle de données ?

Comment construire une base de données en biologie ?

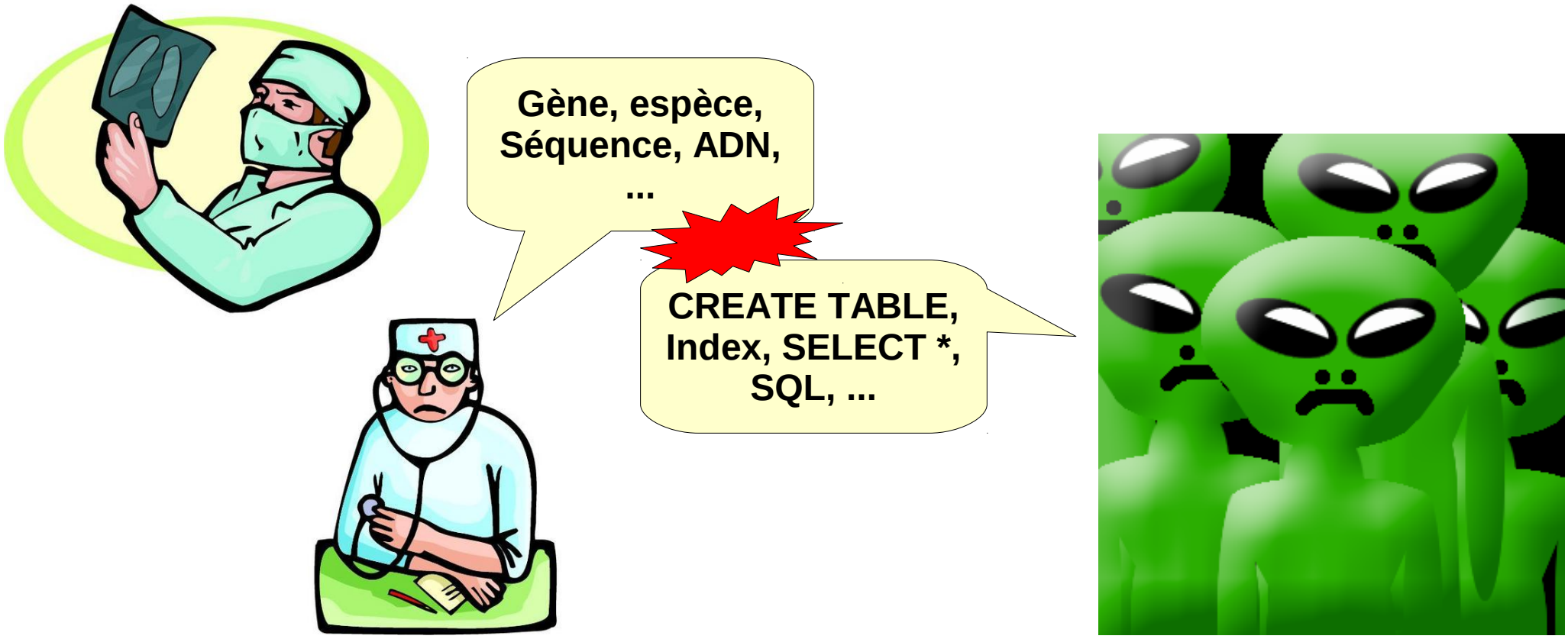
- Comment définir le modèle de données ?
  - On réunit des biologistes...

Comment construire une base de données en biologie ?

- Comment définir le modèle de données ?
  - On réunit des biologistes...
  - … et des informaticiens...

Comment construire une base de données en biologie ?

- Comment définir le modèle de données ?
  - On réunit des biologistes...
  - … et des informaticiens... ...et...



Comment construire une base de données en biologie ?

- **En biologie :**
  - Beaucoup d'exceptions !
  - => Difficile de distinguer les relations 1-1, 1-* / *-1 et *-* !
  - Les informaticiens ne peuvent pas le faire !
  - Exemple :

Table gène

ID	Nom	Séquence
1	Aldostérone synthase	atggcactca...
2	Enveloppe du VIH	tgtacaagaccc...

? est transcrit en ?

Table ARNm

ID	Nom	Séquence
1	Aldostérone synthase	auggcacuca...
2	Enveloppe du VIH	uguacaagaccc...

Identifiant

- Il est important de pouvoir identifier chaque ligne de chaque table de manière unique !
  - Les noms font rarement de bons identifiants
  - Si une ligne a été entrée en double par erreur, il faut pouvoir distinguer les 2 doublons !

Table environnement

Nom	Climat
Forêt	Tropical
Forêt	Equatorial

Table gène

Nom	Séquence
Aldostérone synthase	atggcactca...
Aldostérone synthase	atggcactca...
Aldostérone synthase	atggcactca...

- **=> on ajoute systématiquement un identifiant (=ID)**
  - Généralement numérique (1, 2, 3...)
  - Parfois alphanumérique dans les grandes bases de bioinformatique

Redondance

- Lors de la définition du modèle de données, il faut **éviter au maximum la redondance**
  - Chaque donnée ne doit être présente qu'à un seul endroit
  - Cela facilite les mises à jour

Table personne

Identifiant	Nom	Date de naissance	Taille	Poids	Indice de masse corporelle
1	LAMY	18/06/1979	170	62	21,5
2	X	25/03/1957	165	54	19,8
3	Y	04/01/1982	180	90	27,8

Redondance

- Il y a-t-il des redondances ?

Table gène

Identifiant	Nom	Longueur	Séquence
1	Aldostérone synthase	1512	atggcactcagggcaaaggcaga...
2	Enveloppe du VIH	105	tgtacaagacccaacaacaacacaa...

Redondance

- Il y a-t-il des redondances ?

Table gène

Identifiant	Nom	Longueur	Séquence
1	Aldostérone synthase	1512	atggcactcagggcaaaggcaga...
2	Enveloppe du VIH	105	tgtacaagacccaacaacaacacaa...

Redondance

- Il y a-t-il des redondances ?

Table patient

Identifiant	Nom	Date de naissance
1	LAMY	18/06/1979
2	X	25/03/1957
3	Y	04/01/1982

1 A pour lit 1

Table lit

ID	Etage
1	1
2	1

* Est dans le service 1

ID	Etage	Service
1	1	Cardiologie
2	2	Pneumologie

Table service

Redondance

- Il y a-t-il des redondances ?

Table patient

Identifiant	Nom	Date de naissance
1	LAMY	18/06/1979
2	X	25/03/1957
3	Y	04/01/1982

1 A pour lit 1

Table lit

ID	Etage
1	1
2	1

* Est dans le service 1

ID	Etage	Service
1	1	Cardiologie
2	2	Pneumologie

Table service

Redondance

- Il y a-t-il des redondances ?

Table gène

ID	Nom	Séquence
1	Aldostérone synthase	atggcactca...
2	Enveloppe du VIH	tgtacaagaccc...

1 Transcription *

Table ARNm

ID	Nom	Séquence
1	Aldostérone synthase	auggcacuca...
2	Enveloppe du VIH	uguacaagaccc...

1 Traduction 1

1 Produit la protéine *

Table protéine

ID	Nom	Séquence
1	Aldostérone synthase	VAPIL...
2	Enveloppe du VIH	WRHC...

Redondance

- Il y a-t-il des redondances ?

Table gène

ID	Nom	Séquence
1	Aldostérone synthase	atggcactca...
2	Enveloppe du VIH	tgtacaagaccc...

1 Transcription *

Table ARNm

ID	Nom	Séquence
1	Aldostérone synthase	auggcacuca...
2	Enveloppe du VIH	uguacaagaccc...

1 Traduction 1

1 Produit la protéine *

Table protéine

ID	Nom	Séquence
1	Aldostérone synthase	VAPIL...
2	Enveloppe du VIH	WRHC...

Créer une base de données

- Un serveur MySQL peut gérer plusieurs bases de données
- Pour créer une nouvelle base de données :
  CREATE DATABASE nom_de_la_base;
- Pour utiliser une base de données (nouvellement créée ou non) :
  USE nom_de_la_base;
- Pour supprimer une base de données :
  DROP DATABASE nom_de_la_base;
  - **Attention, toute la base est perdue (modèle et données) !**

Créer des tables en SQL

- Pour créer une table :

```
CREATE TABLE nom_de_la_table (
nom_attribut_1  type_de_donnée,
nom_attribut_2  type_de_donnée,
nom_attribut_3  type_de_donnée,
...
);
```

Types de données

- Les principaux types de données en SQL :
  - INTEGER : nombre entier
  - FLOAT : nombre à virgule (« flottant »)
  - CHAR(n) : chaîne de caractères de n caractères
  - VARCHAR(n) : chaîne de caractères de n caractères ou moins
  - TEXT : chaîne de caractères sans limite de taille
  - DATE : date (ex : 29/01/2010)
  - TIME : heure (ex : 10 h 30, 14 secondes)
  - TIMESTAMP : date + heure

Types de données

- NOT NULL indique un attribut obligatoire
  - Il n'est pas possible de laisser cet attribut vide
- UNIQUE indique que deux lignes ne peuvent pas avoir la même valeur pour cet attribut
- AUTO_INCREMENT indique que la valeur de l'attribut est calculé automatiquement en augmentant de 1 par rapport aux lignes précédentes (à utiliser pour les identifiants !)
- PRIMARY KEY indique une « clef primaire » permettant d'optimiser les recherches ; c'est obligatoire pour les identifiants

Créer des tables en SQL

- Exemple :

```
CREATE TABLE personne (
ID  INTEGER NOT NULL AUTO_INCREMENT
PRIMARY KEY,
nom  VARCHAR(255) NOT NULL,
date_de_naissance  DATE,
taille  INTEGER,
poids  FLOAT);
```

Table personne

Identifiant	Nom	Date de naissance	Taille	Poids
1	LAMY	18/06/1979	170	62
2	X	25/03/1957	165	54,5
3	Y	04/01/1982	180	90

Et les relations ?

- SQL ne permet pas de créer des relations en tant que telles !
- Il faut donc ajouter des attributs correspondant aux relations

Et les relations ?

- Pour les relations **1-1** :
  - On ajoute dans l'une des tables un attribut contenant l'identifiant de l'autre élément

Table patient

ID	Nom	Date de naissance
1	LAMY	18/06/1979
2	X	25/03/1957
3	Y	04/01/1982

Table lit

ID	Service
1	Cardiologie
2	Pneumologie

Et les relations ?

- Pour les relations **1-1** :
  - On ajoute dans l'une des tables un attribut contenant l'identifiant de l'autre élément

Table patient

ID	Nom	Date de naissance	ID lit
1	LAMY	18/06/1979	2
2	X	25/03/1957	0
3	Y	04/01/1982	1

Table lit

ID	Service
1	Cardiologie
2	Pneumologie

Et les relations ?

- Pour les relations **1-1** :
  - On ajoute dans l'une des tables un attribut contenant l'identifiant de l'autre élément

Table patient

ID	Nom	Date de naissance
1	LAMY	18/06/1979
2	X	25/03/1957
3	Y	04/01/1982

Table lit

ID	Service	ID patient
1	Cardiologie	15
2	Pneumologie	1

Et les relations ?

- Pour les relations **1-*** / ***-1** :
  - On ajoute dans la table dont chaque ligne est associée à un seul élément, un attribut contenant l'identifiant de l'autre élément
    - Astuce : il s'agit de la table du côté de l'étoile « * »

Table espèce

ID	Nom	Nom latin
1	Homme	Homo sapiens
2	Poisson zèbre	Danio rerio

Table gène

ID	Nom	Séquence
1	Aldostérone synthase	atggcactca...
2	Enveloppe du VIH	tgtacaagaccc...
3	BMP2	gtccgctaa...

Et les relations ?

- Pour les relations **1-*** / ***-1** :
  - On ajoute dans la table dont chaque ligne est associée à un seul élément, un attribut contenant l'identifiant de l'autre élément
    - Astuce : il s'agit de la table du côté de l'étoile « * »

Table espèce

ID	Nom	Nom latin
1	Homme	Homo sapiens
2	Poisson zèbre	Danio rerio

Table gène

ID	Nom	Séquence	ID espèce
1	Aldostérone synthase	atggcactca...	1
2	Enveloppe du VIH	tgtacaagaccc...	1
3	BMP2	gtccgctaa...	35

Et les relations ?

- Pour les relations ***-*** :
  - Il faut créer une nouvelle table qui assure la liaison
    - Cette table contiendra une ligne par couple

Table espèce

ID	Nom	Nom latin
1	Homme	Homo sapiens
2	Poisson zèbre	Danio rerio

* habite *

Table environnement

ID	Nom
1	Forêt tropicale
2	Mangrove

Et les relations ?

- Pour les relations ***-*** :
  - Il faut créer une nouvelle table qui assure la liaison
    - Cette table contiendra une ligne par couple

Table espèce

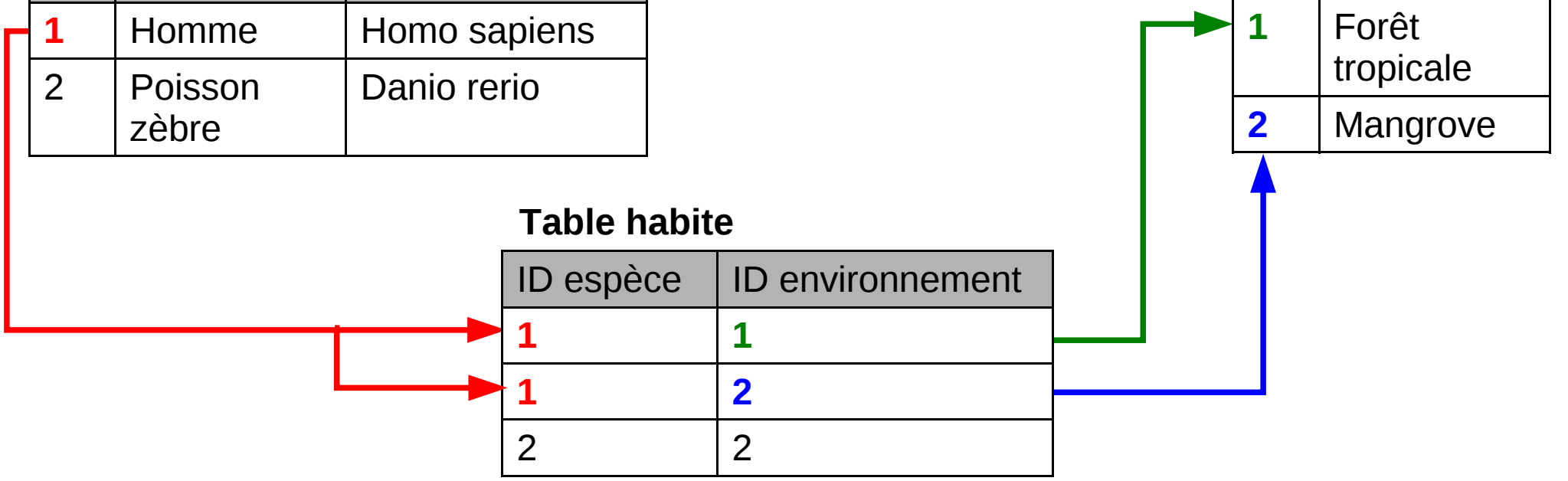
ID	Nom	Nom latin
1	Homme	Homo sapiens
2	Poisson zèbre	Danio rerio

Table environnement

ID	Nom
1	Forêt tropicale
2	Mangrove

Table habite

ID espèce	ID environnement
1	1
1	2
2	2



Et les relations ?

- Exemple :

```
CREATE TABLE patient (
ID  INTEGER NOT NULL AUTO_INCREMENT PRIMARY KEY,
nom  VARCHAR(255) NOT NULL,
date_de_naissance  DATE,
ID_lit INTEGER);

CREATE TABLE lit (
ID  INTEGER NOT NULL AUTO_INCREMENT PRIMARY KEY,
service  VARCHAR(255));
```

Table patient

ID	Nom	Date de naissance	ID lit
1	LAMY	18/06/1979	2
2	X	25/03/1957	51
3	Y	04/01/1982	12

Table lit

ID	Service
1	Cardiologie
2	Pneumologie

Et les relations ?

- Exemple :

```
CREATE TABLE especce (
ID  INTEGER NOT NULL AUTO_INCREMENT PRIMARY KEY,
nom  VARCHAR(255) NOT NULL,
nom_latin  VARCHAR(255) NOT NULL);

CREATE TABLE environnement (
ID  INTEGER NOT NULL AUTO_INCREMENT PRIMARY KEY,
nom  VARCHAR(255) NOT NULL);
```

Table espèce

ID	Nom	Nom latin
1	Homme	Homo sapiens
2	Poisson zèbre	Danio rerio

Table environnement

ID	Nom
1	Forêt tropicale
2	Mangrove

Et les relations ?

- Exemple (suite) :

```
CREATE TABLE habite (
ID_espece  INTEGER NOT NULL,
ID_environnement  INTEGER NOT NULL);
```

Table habite

ID espèce	ID environnement
1	1
1	2
2	2

Supprimer une table

- Pour supprimer une table :
  DROP TABLE nom_de_la_table;
- **Attention :**
  - **Le modèle de la table est perdu !**
  - **Toutes les données de la table sont perdues !**
    - **Y compris les relations si elles étaient dans un attribut de la table**

Insérer des lignes dans une table

- Insertion d'une ligne dans une table :
  INSERT INTO nom_de_la_table
  VALUES (attribut_1, attribut_2, attribut_3,...);
- Exemple :
  INSERT INTO patient
  VALUES (NULL, "LAMY", "1979-06-18", 2);
- NULL => l'ID sera calculé automatiquement
- Attention aux guillemets pour les chaînes de caractères !

Table patient

ID	Nom	Date de naissance	ID lit
1	LAMY	18/06/1979	2
2	X	25/03/1957	51
3	Y	04/01/1982	12

Modifier une table existante

- Pour ajouter une colonne dans une table :
  ALTER TABLE nom_de_la_table
  ADD nouvel_attribut type_de_données;
- Pour supprimer une colonne :
  ALTER TABLE nom_de_la_table
  DROP COLUMN attribut;
- Exemple :
  ALTER TABLE espèce
  ADD menacé INTEGER;

Index

- Pour optimiser les recherches dans la base de données, il est possible de créer des index
  CREATE INDEX nom_de_lindex
  ON table(attribut_1, ...);

- Exemple :
  CREATE INDEX index_des_noms
  ON patient(nom);

- Il est possible d'indexer sur plusieurs attributs pour optimiser des recherches sur plusieurs critères :
  CREATE INDEX index_2
  ON patient(nom, prenom);

Exercice 1

- On souhaite créer une base de données des différentes espèces d'orchidées (plus de 25 000 !) avec MySQL.
- La base de données doit contenir des informations sur les espèces d'orchidées, les genres d'orchidées, et les milieux dans lesquels les orchidées poussent.
- Pour chaque espèce d'orchidée, on souhaite pouvoir indiquer son nom, ainsi que les critères suivants utilisés pour les distinguer : couleur des fleurs, nombre d'étamine, présence de rhizome.
- Les milieux peuvent être humides ou secs, ensoleillés ou non.
- Quels sont les tables dont nous avons besoin ?
  Quels sont leurs attributs ?
  Quelles sont les relations entre les tables ?
- Écrire le code SQL permettant de créer la base de données.

Exercice 1

- L'épipactis des marais est une orchidée du genre épipactis, de couleur pourpre, avec une étamine et un rhizome, qui pousse dans les marais (milieu humide peu ensoleillé)
- Écrire le code SQL permettant d'ajouter l'épipactis des marais dans la base de données

Exercice 2

- Nous souhaitons construire une base de données pour un laboratoire d'analyses médicales. Ce laboratoire emploie plusieurs biologistes, qui effectuent des prélèvements chez des patients, et ensuite analysent ces prélèvements.
- Les biologistes sont identifiés par leur nom et leur prénom, et les patients par leur nom, leur prénom et leur date de naissance.
- Pour chaque analyse, la base doit indiquer le patient, le ou les biologistes, la date d'analyse, le nom de l'analyse et le résultat (sous forme de chiffre).
- Quels sont les tables dont nous avons besoin ?
  Quels sont leurs attributs ?
  Quelles sont les relations entre les tables ?
- Écrire le code SQL permettant de créer la base de données.