

Démonstration à l'aide de la base biblio

- Connexion

Pour se connecter à la base biblio, au prompt linux, taper :

```
<prompt linux>$ psql biblio
biblio=>
```

- Lister les tables

```
biblio=> \d
List of relations
Schema | Name | Type | Owner
-----+-----+-----+-----
public | article_x_auteur | table | bernard
public | articles | table | bernard
public | auteurs | table | bernard
public | journaux | table | bernard
(4 rows)
```

- Voir le schéma d'une table

```
biblio=> \d auteurs
Table "public.auteurs"
Column | Type | Collation | Nullable | Default
-----+-----+-----+-----+-----
nom | text | | not null |
prenom | text | | not null |
Indexes:
    "auteurs_pkey" PRIMARY KEY, btree (nom, prenom)
Referenced by:
    TABLE "article_x_auteur" CONSTRAINT "article_x_auteur_nom_prenom_fkey"
    FOREIGN KEY (nom, prenom) REFERENCES auteurs(nom, prenom)
```

- Requête simple : que contient la table auteurs ?

```
biblio=> SELECT * FROM auteurs ;
 nom | prenom
-----+-----
Lin | Brian Y.
Crick | Francis
Gish | W.
[...]
Lipman | David J.
Margoliash | E.
Tedeschi | A.
(61 rows)
```

N.B. Pas simple de vérifier que chaque auteur est bien unique.

- Choix des attributs (colonnes), renommage

```
biblio=> SELECT abbreviation AS journal, editeur AS "edited by" FROM journaux;
 journal | edited by
-----+-----
Science | American Association for the Advancement of Science
Genetics | Genetics Society of America
J. Comput. Biol. | Mary Ann Liebert
[...]
Amer. J. Hum. Genet. | American Society of Human Genetics
J. Mol. Biol. | Academic Press
Acta Cryst. | International Union of Crystallography
(12 rows)
```

Réponse unique

```
biblio=> SELECT DISTINCT editeur FROM journaux;
          editeur
```

```
-----
Academic Press
Genetics Society of America
Royal Society
[...]
Nature Publishing Group
National Academy of Sciences
International Union of Crystallography
(11 rows)
```

Ordre

```
biblio=> SELECT abbreviation as journal,editeur AS "edited by"
biblio-> FROM journaux ORDER BY "edited by";
          journal | edited by
```

```
-----+-----
J. Mol. Biol. | Academic Press
Science | American Association for the Advancement of ...
Amer. J. Hum. Genet. | American Society of Human Genetics
[...]
Nucl. Acids Res. | Oxford University Press
Bioinformatics | Oxford University Press
Phil. Trans. R. Soc. Lon. B | Royal Society
(12 rows)
```

- Opérations, fonctions

Arithmétique

```
biblio=> SELECT 3 + 4 ;
          ?column?
-----
          7
(1 row)
```

Plusieurs opérations

```
biblio=> SELECT 3 * 4 AS produit, 3 + 4 AS somme ;
          produit | somme
-----+-----
          12 | 7
(1 row)
```

Fonction sur les chaînes

```
biblio=> SELECT SUBSTRING('frugalité',6,3);
          SUBSTRING
-----
          lit
(1 row)
```

```
biblio=> SELECT SUBSTRING(titre,1,30) AS "titre tronqué" FROM articles;
          titre tronqué
```

```
-----
Archaeal-eubacterial mergers i
A tool for multiple sequence a
On the statistical significanc
[...]
DNA sequencing with chain-term
The interpretation of diffuse
A cytological and genetical st
(25 rows)
```

Choix des entités (lignes), expression booléenne

```
biblio=> SELECT SUBSTRING(titre,1,30) AS "titre tronqué", annee, volume
```

```
biblio=> FROM articles WHERE annee >= 1990;
```

titre tronqué	annee	volume
Archaeal-eubacterial mergers i	1996	93
Basic local alignment search t	1990	215
On the sequencing and assembly	2002	99
[...]		
An overview on the distributio	2000	7
Oases: robust de novo RNA-seq	2012	28
Identifying potential tRNA gen	1991	220

(8 rows)

Calcul sur les attributs.

```
biblio=> SELECT SUBSTRING(titre,1,30) AS "titre tronqué", annee, debut, fin
```

```
biblio=> FROM articles WHERE fin - debut <= 9;
```

titre tronqué	annee	debut	fin
Archaeal-eubacterial mergers i	1996	1071	1076
A tool for multiple sequence a	1989	4412	4415
Basic local alignment search t	1990	403	410
[...]			
Oases: robust de novo RNA-seq	2012	1086	1092
Molecular structure of nucleic	1953	737	738
DNA sequencing with chain-term	1977	5463	5467

(13 rows)

Comparaison de chaînes

```
biblio=> SELECT nom_journal FROM journaux WHERE nom_journal LIKE 'Jour%';
```

```
nom_journal
```

Journal of Computational Biology
Journal of Molecular Biology

(2 rows)

```
biblio=# SELECT nom_journal FROM journaux WHERE nom_journal ~ '^Jour';
```

```
nom_journal
```

Journal of Computational Biology
Journal of Molecular Biology

(2 rows)

```
biblio=> SELECT nom_journal FROM journaux WHERE nom_journal LIKE '%Sci%';
```

```
nom_journal
```

Science
Proceedings of the National Academy of Sciences of the United States of America
Philosophical Transactions of the Royal Society of London B: Biological Sciences

(3 rows)

```
biblio=> SELECT nom_journal FROM journaux WHERE nom_journal ILIKE '%sta%';
```

```
nom_journal
```

Proceedings of the National Academy of Sciences of the United States of America
Acta Crystallographica

(2 rows)

Expression booléenne

```
biblio=> SELECT SUBSTRING(titre,1,30) as "titre tronqué",debut,fin,annee
biblio->      FROM articles WHERE annee >= 1990 AND fin - debut <= 9;
```

titre tronqué	debut	fin	annee
Archaeal-eubacterial mergers i	1071	1076	1996
Basic local alignment search t	403	410	1990
On the sequencing and assembly	4145	4146	2002
Structural insights into a yea	765	772	2005
An overview on the distributio	193	201	2000
Oases: robust de novo RNA-seq	1086	1092	2012

(6 rows)

- Jointures

Jointure croisée (= produit cartésien)

```
biblio=> SELECT nom, editeur FROM auteurs CROSS JOIN journaux;
biblio->      nom | editeur
```

nom	editeur
Lin	American Association for the Advancement of Science
Crick	American Association for the Advancement of Science
Gish	American Association for the Advancement of Science
[...]	
Lipman	American Association for the Advancement of Science
Margoliash	American Association for the Advancement of Science
Tedeschi	American Association for the Advancement of Science
Lin	Genetics Society of America
Crick	Genetics Society of America
Gish	Genetics Society of America
[...]	
Lipman	Genetics Society of America
Margoliash	Genetics Society of America
Tedeschi	Genetics Society of America
[...]	
Lin	International Union of Crystallography
Crick	International Union of Crystallography
Gish	International Union of Crystallography
[...]	
Lipman	International Union of Crystallography
Margoliash	International Union of Crystallography
Tedeschi	International Union of Crystallography

(732 rows)

Pourquoi 732 ?

```
biblio=> SELECT 61 * 12;
biblio->      ?column?
```

732

(1 row)

Autre exemple

```
biblio=> SELECT SUBSTRING(titre,1,40) as "titre tronqué",abreviation
biblio->      FROM articles CROSS JOIN journaux;
```

titre tronqué	abreviation
Archaeal-eubacterial mergers in the orig	Science
Archaeal-eubacterial mergers in the orig	Genetics
Archaeal-eubacterial mergers in the orig	J. Comput. Biol.
[...]	
A cytological and genetical study of tri	Amer. J. Hum. Genet.
A cytological and genetical study of tri	J. Mol. Biol.
A cytological and genetical study of tri	Acta Cryst.

(300 rows)

Sélection sur un attribut commun

```
biblio=> SELECT SUBSTRING(titre,1,40) as "titre tronqué",abreviation
```

```
biblio-> FROM articles CROSS JOIN journaux
```

```
biblio-> WHERE articles.issn = journaux.issn;
```

titre tronqué	abreviation
Archaeal-eubacterial mergers in the orig	Proc. Natl. Acad. Sci. USA
A tool for multiple sequence alignment	Proc. Natl. Acad. Sci. USA
On the statistical significance of nucle	Nucl. Acids Res.
[...]	
DNA sequencing with chain-terminating in	Proc. Natl. Acad. Sci. USA
The interpretation of diffuse X-ray diag	Acta Cryst.
A cytological and genetical study of tri	Genetics

(25 rows)

- Jointure interne : 2 formulations, la seconde ne joint que sur égalité, colonnes implicites

```
biblio=> SELECT SUBSTRING(titre,1,30) as "titre tronqué",articles.issn,
```

```
biblio-> journaux.issn,abreviation
```

```
biblio-> FROM articles JOIN journaux ON articles.issn = journaux.issn;
```

titre tronqué	issn	issn	abreviation
Archaeal-eubacterial mergers i	0027-8424	0027-8424	Proc. Natl. Acad. Sci. USA
A tool for multiple sequence a	0027-8424	0027-8424	Proc. Natl. Acad. Sci. USA
On the statistical significanc	0305-1048	0305-1048	Nucl. Acids Res.
[...]			
DNA sequencing with chain-term	0027-8424	0027-8424	Proc. Natl. Acad. Sci. USA
The interpretation of diffuse	2059-7983	2059-7983	Acta Cryst.
A cytological and genetical st	0016-6731	0016-6731	Genetics

(25 rows)

```
biblio=> SELECT SUBSTRING(titre,1,30) as "titre tronqué",articles.issn,
```

```
biblio-> journaux.issn,abreviation
```

```
biblio-> FROM articles NATURAL JOIN journaux;
```

titre tronqué	issn	issn	abreviation
Archaeal-eubacterial mergers i	0027-8424	0027-8424	Proc. Natl. Acad. Sci. USA
A tool for multiple sequence a	0027-8424	0027-8424	Proc. Natl. Acad. Sci. USA
On the statistical significanc	0305-1048	0305-1048	Nucl. Acids Res.
[...]			
DNA sequencing with chain-term	0027-8424	0027-8424	Proc. Natl. Acad. Sci. USA
The interpretation of diffuse	2059-7983	2059-7983	Acta Cryst.
A cytological and genetical st	0016-6731	0016-6731	Genetics

(25 rows)

Autre exemple

```
biblio=> SELECT nom,prenom,titre FROM article_x_auteur NATURAL JOIN articles;
```

nom	prenom	titre
Fitch	W. M.	Construction of phylogenetic trees
Schbath	Sophie	An overview on the distribution of word counts in ...
Creighton	Harriet B.	A correlation of cytological and genetical crossing-over ...
[...]		
Miadonnaqi	A.	Human basophil degranulation triggered by very dilute ...
Fox	George E.	Phylogenetic structure of the prokaryotic domain: The ...
Kececioglu	J. D.	A tool for multiple sequence alignment

(69 rows)

- Pour grouper par article et ordonner les auteurs

```
biblio=> SELECT nom,prenom,titre FROM article_x_auteur NATURAL JOIN articles
biblio-> ORDER BY titre,rang;
```

nom	pre nom	titre
Creighton	Harriet B.	A correlation of cytological and genetical crossing-over ...
McClintock	Barbara	A correlation of cytological and genetical crossing-over ...
McClintock	Barbara	A cytological and genetical study of triploid maize
[...]		
Lin	Brian Y.	tRNAscan-SE 2.0: improved detection and functional ...
Mak	Allysia J.	tRNAscan-SE 2.0: improved detection and functional ...
Lowe	Todd M.	tRNAscan-SE 2.0: improved detection and functional ...

Double jointure naturelle

```
biblio=> SELECT nom,prenom,SUBSTRING(titre,1,20) AS "titre tronqué",abreviation
biblio-> FROM article_x_auteur NATURAL JOIN articles
```

```
biblio-> NATURAL JOIN journaux;
```

nom	pre nom	titre tronqué	abreviation
Fitch	W. M.	Construction of phyl	Science
Schbath	Sophie	An overview on the d	J. Comput. Biol.
Creighton	Harriet B.	A correlation of cyt	Proc. Natl. Acad. Sci. USA
[...]			
Miadonnai	A.	Human basophil degra	Nature
Fox	George E.	Phylogenetic structu	Proc. Natl. Acad. Sci. USA
Kececioglu	J. D.	A tool for multiple	Proc. Natl. Acad. Sci. USA

(69 rows)

Exercice : Qui a co-publié avec David J. Lipman ?

Une solution, créée en trois étapes :

1. Toutes les paires d'auteurs (y compris X,X) pour chaque article :

```
biblio=> SELECT aa1.* , aa2.* FROM article_x_auteur AS aa1
biblio-> JOIN article_x_auteur AS aa2 ON aa1.doi = aa2.doi;
→ 337 rows
```

2. Toutes les paires d'auteurs non redondantes pour chaque article (seulement les noms) :

```
biblio=> SELECT aa1.nom || ', ' || aa1.prenom AS auteur1,
biblio-> aa2.nom || ', ' || aa2.prenom AS auteur2
biblio-> FROM article_x_auteur AS aa1 JOIN article_x_auteur AS aa2
biblio-> ON aa1.doi = aa2.doi
biblio-> AND NOT (aa1.nom = aa2.nom AND aa1.prenom = aa2.prenom);
→ 268 rows
```

3. Les auteurs qui ont au moins un article en commun avec Lipman (lui-même non compris) :

```
biblio=> SELECT DISTINCT aa2.nom || ', ' || aa2.prenom AS "co-auteur"
biblio-> FROM article_x_auteur AS aa1 JOIN article_x_auteur AS aa2
biblio-> ON aa1.doi=aa2.doi
biblio-> AND NOT (aa1.nom = aa2.nom AND aa1.prenom = aa2.prenom)
biblio-> WHERE aa1.nom = 'Lipman' ORDER BY "co-auteur";
```

co-auteur
Altschul, S. F.
Gish, W.
Kececioglu, J. D.
Miller, W.
Myers, Eugene W.
Smith, T. F.
Waterman, M. S.
Wilbur, W. J.

(8 rows)

- Groupes

```
biblio=> SELECT editeur FROM journaux;
          editeur
-----
American Association for the Advancement of Science
Genetics Society of America
Mary Ann Liebert
[...]
American Society of Human Genetics
Academic Press
International Union of Crystallography
(12 rows)

biblio=> SELECT editeur FROM journaux GROUP BY editeur;
          editeur
-----
Academic Press
Genetics Society of America
Royal Society
[...]
Nature Publishing Group
National Academy of Sciences
International Union of Crystallography
(11 rows)
```

- Aggrégats

```
biblio=> SELECT count(editeur) FROM journaux;
count
-----
12
(1 row)

biblio=> SELECT count(editeur) FROM journaux WHERE editeur LIKE '%Press%';
count
-----
3
(1 row)

biblio=> SELECT editeur,count(editeur) FROM journaux GROUP BY editeur;
          editeur                                     | count
-----+-----
Academic Press                                     |      1
Genetics Society of America                       |      1
Royal Society                                     |      1
[...]
Nature Publishing Group                           |      1
National Academy of Sciences                      |      1
International Union of Crystallography            |      1
(11 rows)

biblio=> SELECT editeur,count(editeur) FROM journaux
biblio->      WHERE editeur LIKE '%Press%' GROUP BY editeur;
          editeur                                     | count
-----+-----
Academic Press                                     |      1
Oxford University Press                           |      2
(2 rows)
```

```
biblio=> SELECT editeur,count(editeur) FROM journaux
biblio->      WHERE editeur LIKE '%Press%'
biblio->      GROUP BY editeur HAVING count(editeur) > 1;
```

editeur	count
Oxford University Press	2

(1 row)

```
biblio=> SELECT nom,prenom,count((nom,prenom)) FROM article_x_auteur
biblio->      GROUP BY (nom,prenom) HAVING count((nom,prenom)) > 1;
```

nom	prenom	count
Altschul	S. F.	2
Lindquist	Susan L.	2
Lipman	David J.	3
Myers	Eugene W.	2
Smith	T. F.	2
McClintock	Barbara	2
Waterman	M. S.	2

(7 rows)

```
biblio=> SELECT nom,prenom FROM article_x_auteur
biblio->      GROUP BY (nom,prenom) HAVING count((nom,prenom)) > 1;
```

nom	prenom
Altschul	S. F.
Lindquist	Susan L.
Lipman	David J.
Myers	Eugene W.
Smith	T. F.
McClintock	Barbara
Waterman	M. S.

(7 rows)