MAPPI

5 tâches,

- *Tâche 1 : Nouvelles structures d'index pour la recherche de motifs approchés*
- *Tâche 2 :* Mapping *pour la métagenomique et la métatranscriptomique*
- *Tâche 3 : Outils d'assemblage pour les* NGS
- *Tâche 4 : Assemblage guidé de données de métatranscriptomique*
- *Tâche 5 : Pipeline bioinformatique*

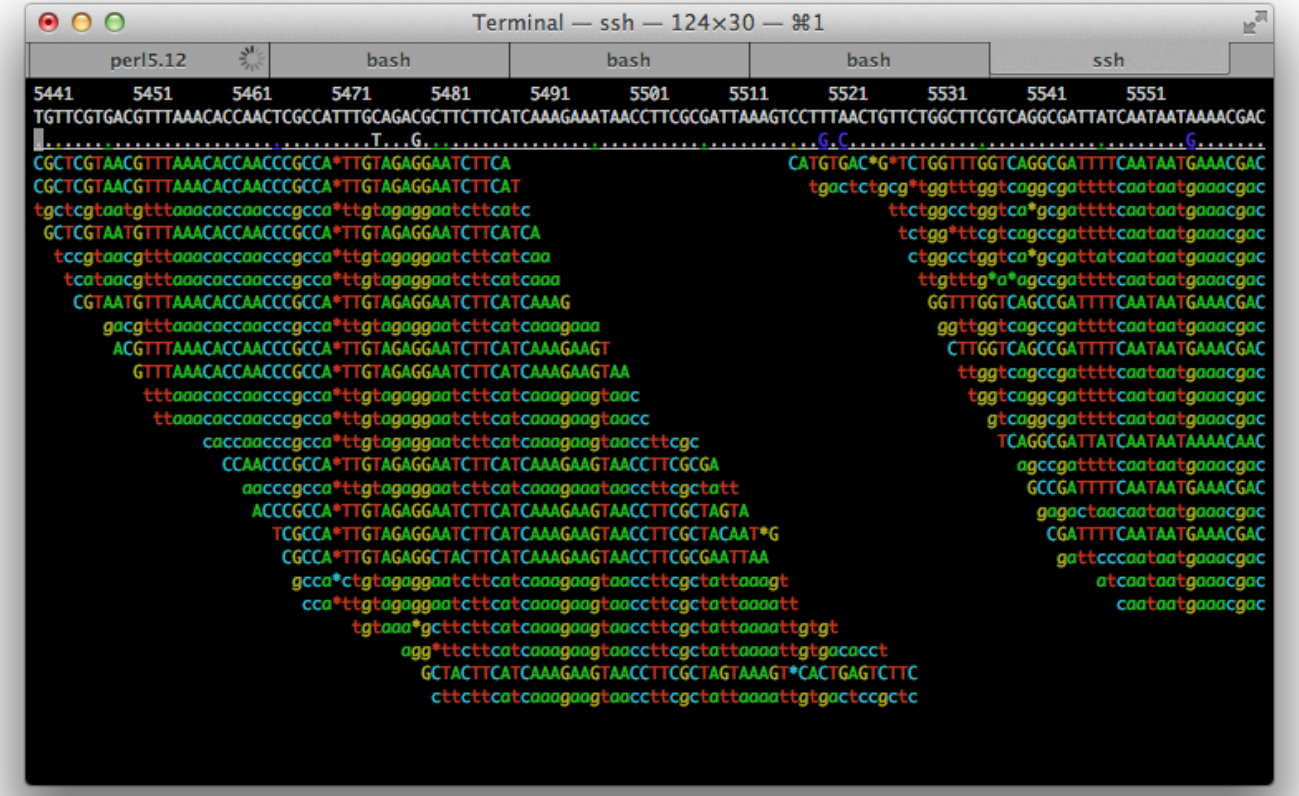
MAPPI

5 tâches, celles que je vais décrire dans le contexte Lillois

- *Tâche 1 : Nouvelles structures d'index pour la recherche de motifs approchés*
- *Tâche 2 :* Mapping *pour la métagenomique et la métatranscriptomique*
- *Tâche 3 : Outils d'assemblage pour les* NGS
- *Tâche 4 : Assemblage guidé de données de métatranscriptomique*
- *Tâche 5 : Pipeline bioinformatique*

Tâche 1 : Nouvelles structures d'index pour la recherche de motifs approchés

Contexte : *Read Mapping*



Tâche 1 : Nouvelles structures d'index pour la recherche de motifs approchés

Contexte : *Read Mapping*

Tâche 1 : Nouvelles structures d'index pour la recherche de motifs approchés

Contexte : *Read Mapping*

Réalisé :

1. Portage de l'algorithme de Wu-Mamber sur GPU

   [Bit-Parallel Multiple Pattern Matching. T. T. Tran, M. Giraud, J.-S. Varré PPAM / PBC 2011.]

2. Indexation des voisinages des k-mers
   But : profiter de l'efficacité du cache GPU/Processeur
   Deux méthodes d'indexation envisagées :
   - indexation directe (tri des mots $\rightarrow$ recherche dichotomique)
   - hachage parfait

   + non encore publié mais des résultats :
   - mise en oeuvre en OpenCL (fonctionnelle sur CPU et GPU)
   - gain en performance entre x10 et x60
   - prototype de readmapper en cours

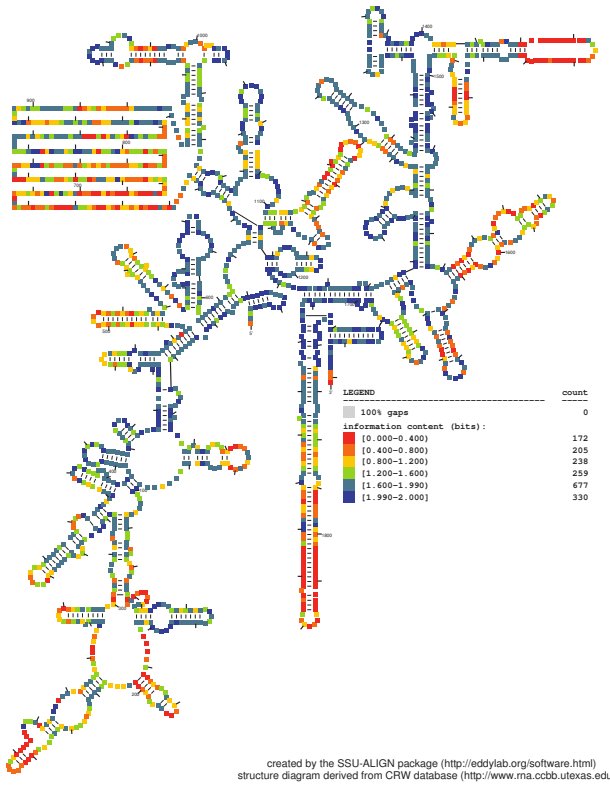
[LIFL] *Tuan Tu Tran*, *Mathieu Giraud*, *Jean-Stéphane Varré* [LIAFA] *Djamal Belazzougui*, *Mathieu Raffinot*

Tâche 4 : Assemblage guidé de données de métatranscriptomique

Contexte : identification d'ARN ribosomiques (16S/18S,23S/28S...)

Buts :
- élimination
- classification

? : Problème nouveau sur données de *métatranscriptomique*

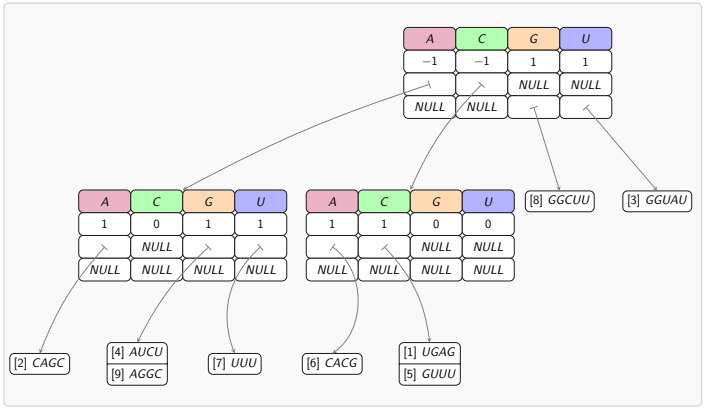


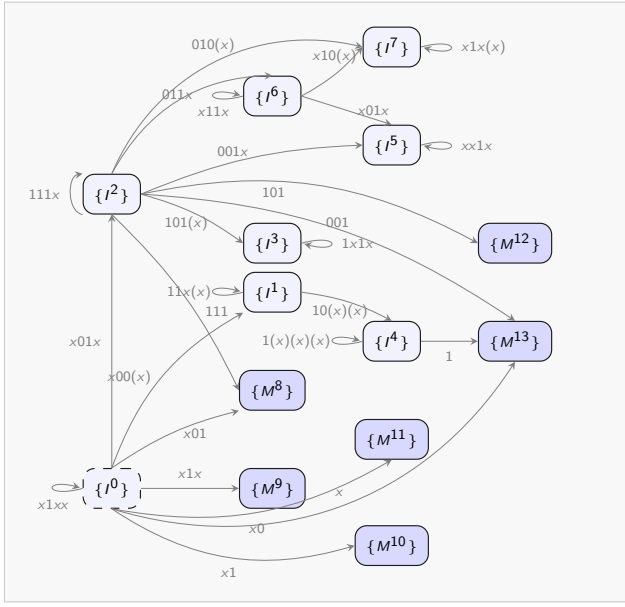
Tâche 4 : Assemblage guidé de données de métatranscriptomique

Contexte : identification d'ARN ribosomiques

Réalisé :

- conception d'un filtre efficace pour la sélection des familles d'ARNr (`SortMeRNA`)
- travail basé sur le *Burst Trie* et *l'automate de Levenstein*





Tâche 4 : Assemblage guidé de données de métatranscriptomique

Contexte : identification d'ARN ribosomiques

En cours :

- communication aux *London Stringology Days*
- publication en cours de soumission
- séjour prévue au Génoscope pour la transition *fin Tâche 4 / début Tâche 2*, 10-13 avril

[LIFL] *Evguenia Kopylova* (ANR), *Laurent Noé*, *Hélène Touzet* [GENOSCOPE] *Olivier Jaillon*