

# Rapport de stage de M2

Nicolas JEANNE

11 mars 2015

# Introduction

En 1982, la découverte par Higgins et al. de nouveaux éléments génétiques communs dans les régions intercistroniques des opérons de *Escherichia coli* et *Salmonella typhimurium* a constitué le premier pas de la recherche sur les Repeated Extragenic Palindrome (REP) (Higgins et al. 1982). En 1991, Gilson et al. ont mis en évidence l'organisation en clusters de ces REP (Gilson et al. 1991), ces clusters ont été caractérisés comme Bacterial Interspersed Mosaic Element (BIME). Chez *E. coli* en 1994, Bachelier et al. ont réussi à catégoriser les REP constituant les BIME en 2 types Y et Z, constituants 3 motifs Y, Z<sup>1</sup>, Z<sup>2</sup>.

Les REP constituent une part non négligeable du génome bactérien, chez *E. coli* K12 ou *S. typhimurium* elles représentent environ 1% de celui-ci (Gilson et al. 1991). Nous les retrouvons chez de nombreux règnes bactériens, notamment chez les pathogènes humains tels que *Escherichia coli*, *Salmonella enterica*, *Neisseria meningitidis*, *Mycobacterium tuberculosis* et *Pseudomonas aeruginosa* mais également chez des pathogènes des plantes comme *Agrobacterium tumefaciens* ou chez des bactéries ubiquitaire, *Deinococcus radiodurans* ou *Pseudomonas putida* par exemple. Les travaux précédents de l'équipe ont permis l'annotation des REP au sein du génome d'*E. coli* et de mettre en évidence le lien existant entre la prolifération des REP et le gène *tnpA<sub>REP</sub>* (Bosc 2014; Weyder 2013), ainsi que la reconstruction des états ancêtres des REP (Bosc 2014). Le rôle exact des REP n'est pas clairement défini, des hypothèses sont avancées sur leur implication dans la régulation de l'expression des gènes, que ce soit en tant que terminateur ou comme site de reconnaissance des enzymes impliquées dans les mécanismes de la transcription.

## Structure des REP et organisations en BIME

La taille des REP varie de 20 à 40 nucléotides, la classification Y,  $Z^1$ ,  $Z^2$  est basée à la fois sur la taille de la séquence consensus de la REP ainsi que sur sa structure secondaire. Par convention, une REP en orientation inversée est nommée iREP (inversed REP) (Ton-Hoang et al. 2012). Un tétra-nucléotide caractéristique de séquence GTAC est présent à l'extrémité 5' des REP, sa séquence est CTAC en 3' pour les iREP. Les différentes classes de REP partagent des nucléotides conservés (Figure 1A). La structure secondaire des REP est caractérisée par sa forme en tige-boucle, le caractère palindromique permet la formation de la tige malgré un mésappariement situé dans la partie centrale de celle-ci (Figure 1B). Pour le génome d'E. coli K12, 93 REP ont été répertoriées comme étant uniques sur les 605 annotées par le laboratoire, les autres sont organisées par paires sous forme de BIME. Une classification a été adoptée comportant 3 entrées, les BIME-1 composées de REP  $Z^1$  et Y apparaissant en paires uniques et dans lesquelles la REP et l'iREP sont séparées par une séquence longue (L). Les BIME-2 constituées de  $Z^2$  et de Y, apparaissant en copies multiples de cette paire dont la REP et l'iREP sont séparées par une séquence courte (S) et une des trois séquences flanquantes (s, l ou r). La troisième catégorie est constituée des BIME dites atypiques qui sont des chimères de BIME-1 et BIME-2, comportant différentes combinaisons de Y,  $Z^1$ ,  $Z^2$ , S, L, s, l et r. Tout comme les BIME-2, nous les retrouvons sous forme de copies multiples (Figure 1C).

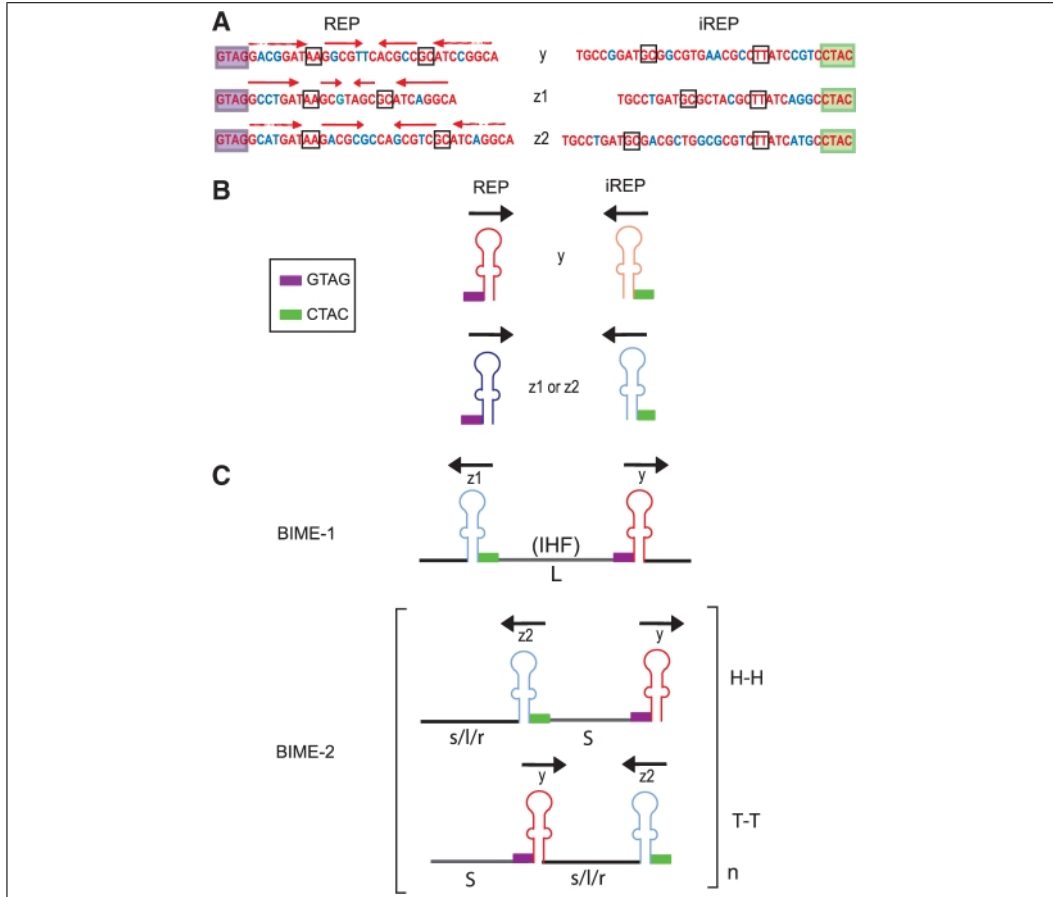


FIGURE 1 – **REP et BIME chez Escherichia coli.** (A) Séquences consensus Y, Z<sup>1</sup> et Z<sup>2</sup> des REP. Le tétra-nucléotide conservé GTAC est encadré en violet, le complémentaire conservé CTAC est encadré en vert, les flèches rouges situent les zones d'appariement de la tige et les positions encadrées en noir sont les zones de mésappariement. Les positions conservées parmi les classes de REP sont en rouge, les positions variables en bleu. (B) Structure secondaire des REP. Les rectangles violets et verts représentent respectivement les tétra-nucléotides conservés GTAC pour les REP et CTAC pour les iREP. Les flèches noires indiquent l'orientation des REP. (C) Structures des BIME-1 et BIME-2. Les BIME-1 sont composées de REP et de iREP Y et Z<sup>1</sup> séparées par une séquence longue (L), les BIME-2 sont composées de Y et Z<sup>2</sup>, de séquences séparatrices courtes (S) et de séquences séparatrices s, l ou r. H-H et T-T dénotent respectivement une organisation tête à tête et queue à queue des REP. (Ton-Hoang et al. 2012).

## Matériel & Méthodes

# Bibliographie

- Julie Bosc. Etude de la dynamique des éléments palindromique répétées ( REP ) chez l ' espèce *Escherichia coli* par une méthode de reconstruction des états ancêtres . Technical report, 2014.
- E Gilson, W Saurin, D Perrin, S Bachellier, and M Hofnung. Palindromic units are part of a new bacterial interspersed mosaic element (BIME). *Nucleic acids research*, 19(7) :1375–1383, 1991. ISSN 03051048.
- Christopher F. Higgins, Giovanna Ferro-Luzzi Ames, Wayne M. Barnes, Jean Marie Clement, and Maurice Hofnung. A novel intercistronic regulatory element of prokaryotic operons. *Nature*, 298(5876) :760–762, August 1982. ISSN 0028-0836. doi : 10.1038/298760a0.
- Bao Ton-Hoang, Patricia Siguier, Yves Quentin, Séverine Onillon, Brigitte Marty, Gwennaele Fichant, and Mick Chandler. Structuring the bacterial genome : Y1-transposases associated with REP-BIME sequences. *Nucleic acids research*, 40(8) :3596–609, April 2012. ISSN 1362-4962. doi : 10.1093/nar/gkr1198. URL <http://nar.oxfordjournals.org/content/40/8/3596.long>.
- Mathias Weyder. Étude de la dynamique de la prolifération des éléments REP chez *Escherichia* et *Shigella* par une approche bioinformatique. Technical report, 2013.