

### **Module : Génétique des Procaryotes**

## **Chapitre IV: les Transposons et les Intégrons**

### **I) Les éléments transposables**

Très schématiquement, on peut considérer que le génome d'une bactérie est constitué par:

- 1) **Le DNA chromosomique:** dans lequel sont codifiées toutes les informations indispensables à la survie et à la multiplication bactériennes.
- 2) **Le DNA extra-chromosomique:** (plasmides, prophages) dans lequel sont codifiés un certain nombre d'informations non indispensables dans les conditions normales mais qui peuvent le devenir dans des situations particulières (ex.: résistance aux antibiotiques).

Cette information génétique contenue dans les plasmides est potentiellement importante: elle doit donc pouvoir circuler rapidement entre bactéries; ceci peut se faire à un premier niveau par les transferts génétiques inter-bactériens. Il arrive cependant qu'un plasmide qui pénètre dans une bactérie réceptrice ne puisse pas se répliquer (enzymes de restriction, problèmes de réplication) ou qu'il ne puisse pas se re-combiner suivant le système classique par homologie. Pour éviter que les gènes importants ne soient alors perdus pour la bactérie réceptrice, l'évolution a mis au point un système de transfert génétique intra-bactérien basé sur les éléments transposables.

#### **1- Définition**

Les chromosomes de bactéries, de virus et de cellules eucaryotes contiennent des morceaux d'ADN qui se déplacent le long du génome. Ce déplacement est appelé **transposition**. Les séquences d'ADN porteuses des gènes nécessaires à ce processus et donc capables de se mouvoir le long des chromosomes sont des **éléments transposables** ou **transposons**. Contrairement à d'autres mécanismes qui réorganisent l'ADN, la transposition ne requiert pas des zones étendues d'homologie entre le transposon et son site de destination (recombinaison illégitime).

Les éléments transposables furent découverts par Barbara MCCLINTOCK en 1951 au cours de ses études de la génétique du maïs, une découverte qui lui valut le prix Nobel en 1983.

#### **2- Les différents types des éléments transposables**

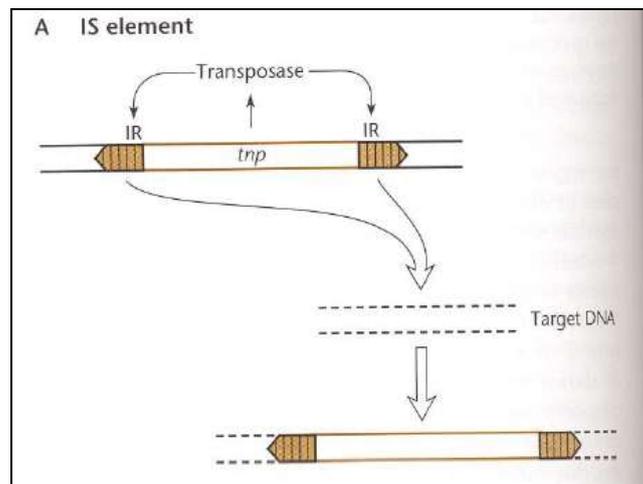
Les éléments transposables peuvent être différenciés en:

**1) Les séquences d'insertion ou IS:** ce sont de petits éléments génétiques qui ne codent que pour l'information nécessaire à leur transposition.

Elles contiennent de 700 à 1600 bp, encadrées aux extrémités par deux courts segments inverses de 15 à 25 bp (IR). Un élément *IS* contient le signal nécessaire à la transposition, qui est catalysée par une transposase.

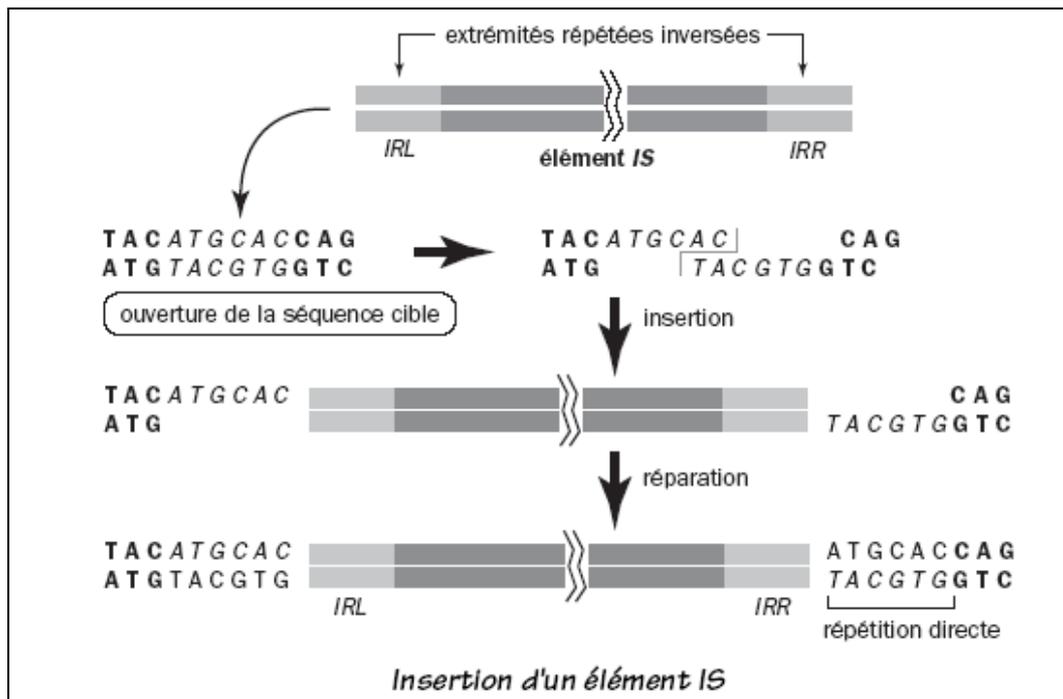
Un élément IS est flanqué aux deux extrémités par des séquences nucléotidiques identiques ou très similaires en sens inverses ou séquence IR (inverted repeats), varient parmi les un élément IS de manière que chaque type IS ait des séquences répétitives inverses propres et caractéristiques.

Entre les séquences répétitives inverses, on trouve un gène qui code pour un enzyme appelé **transposase**, enzyme nécessaire à la transposition qui reconnaît les extrémités des IS avec une grande précision. Chaque élément IS est désigné par le préfixe IS suivi d'un numéro. Chez *E.coli* différents IS sont observés (de IS1 à IS5), avec une longueur de 768 à 1428pb, une séquence répétitive inverse comprise entre 16 et 41 pb, une cible de 3 à 12 pb et un nombre de copies dans le chromosome allant de 1-2 à 10-11.



**Figure 1: l'élément IS**

L'enzyme se lie aux deux extrémités repesées inversées (ou *IR*), reconnaît la séquence cible et y effectue l'insertion. La transposase agit sur la cible en y pratiquant une ouverture formant des "bouts collants" à la manière de ce que ferait une nucléase de restriction. Son gène occupe presque toute la longueur intermédiaire de l'élément, et son promoteur se situe en partie dans l'un des segments répètes inverses, qui est désigné par convention comme extrémité gauche *IRL* (inverted repeat, left). L'autre extrémité est *IRR* (inverted repeat, right).



**Figure 2: Insertion d'un élément IS**

On voit que la petite séquence servant de cible à l'insertion subit **une duplication en tandem** et l'*IS* s'installe entre les deux copies.

La présence des *IS* dans un génome (comme celle des transposons) est responsable d'une proportion importante des mutations chez les bactéries, et constitue un facteur d'évolution considérable. Ces *IS*:

- a) **ont un caractère mutagène** (le gène dans lequel un élément transposable s'est inséré perd généralement sa fonction).
- b) **peuvent jouer un rôle dans l'expression de l'information d'un gène adjacent** (certains *IS* portent un site promoteur).
- c) **semblent jouer un rôle dans l'organisation, l'arrangement de certains gènes.**

**2) Les transposons composites:** qui portent des déterminants autres que ceux nécessaires à leur transposition (résistance aux antibiotiques, production de toxines, dégradation du lactose).

Ces transposons ont une région centrale contenant les gènes supplémentaires flanqués de part et d'autre par des éléments *IS* dont les séquences sont identiques ou très similaires. Les éléments *IS* peuvent être soit dans la même orientation (répétition directe) soit dans l'orientation opposée (répétition inversée). Les noms des transposons composites commencent par le préfixe **Tn**. Par exemple

**Tn 5:** taille (5700pb), marqueurs (Kan<sup>r</sup>), extrémités (IS 50, inversées), la taille de la séquence cible (9pb).

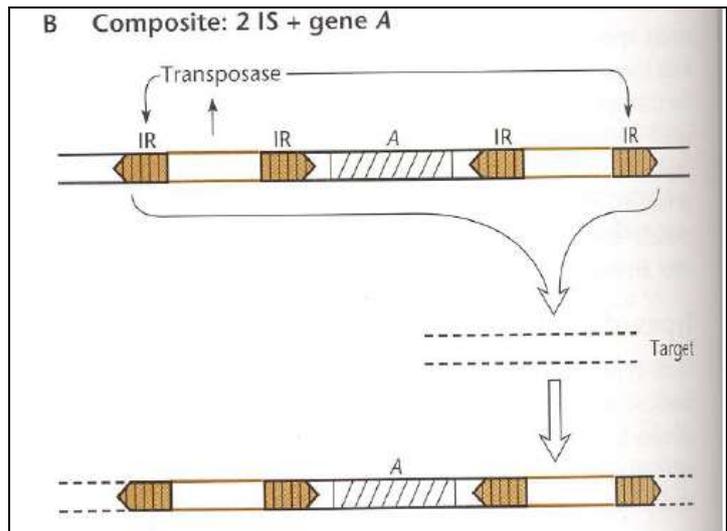


Figure 3: Transposon composite

### 3) Les transposons de la famille Tn3

Ces éléments sont constitués de deux séquences IR encadrant un gène *tnpA*: gène codant pour la transposase, un gène codant pour la résolvasse *tnpR*, un site de résolution interne (SRI) et un gène *bla*, de résistance au  $\beta$ -lactamines. Ils sont très nombreux chez les bactéries Gram négative: Tn3 (9,4 Kb, ApR), Tn4 (9Kb, ApR, SmR, SuR). Chez les Gram positive: Tn551 (5,3Kb, EmR)

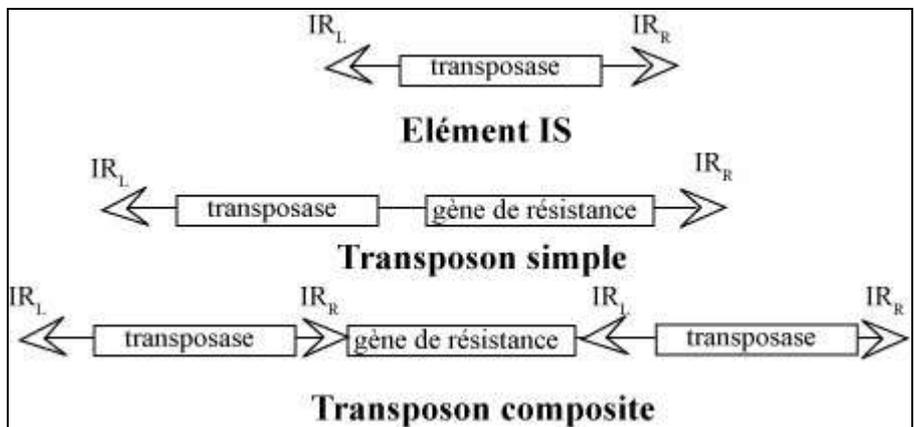


Figure 4: Les différents types de transposons

**4) Transposons conjugatifs:** ces éléments codent pour le gène *xis-Tn* codant pour l'excisionase, le gène *int-Tn* codant pour l'intégrase, le gène *tet-M* codant pour la résistance à la tétracycline, les gènes codant pour les fonctions *tra+* nécessaires au transfert conjugatif. Ces transposons sont détectés chez les Gram+ comme par exemple Tn 916 (16Kb, TcR). Le spectre d'hôte des transposons conjugatifs est large, vers la quasi-totalité des bactéries à Gram+, mais également vers les bactéries Gram-.

### 3) La transposition

Il existe deux types de transposition:

- a) **Transposition conservatrice ou insertion simple:** Une séquence d'ADN est transférée d'un site à un autre, entre un site donneur et un site accepteur.
- b) **Transposition répllicative :** l'élément transposable est transféré d'un site à un autre, tout en restant au site original. Cela conduit à une augmentation du nombre de copies de l'élément transposable.

Selon les éléments transposables, un mode ou l'autre, ou les deux seront employés.

Le premier mode de transposition se fait généralement selon un mode de "couper-coller", on observe généralement lorsqu'il s'agit d'un changement de site sur le même chromosome. Alors que le deuxième type de transposition se fait selon un mode de "copier-coller", le plus souvent il est observé lorsqu'il s'agit d'une transposition plasmide-plasmide.

### 4) Mécanisme de la transposition

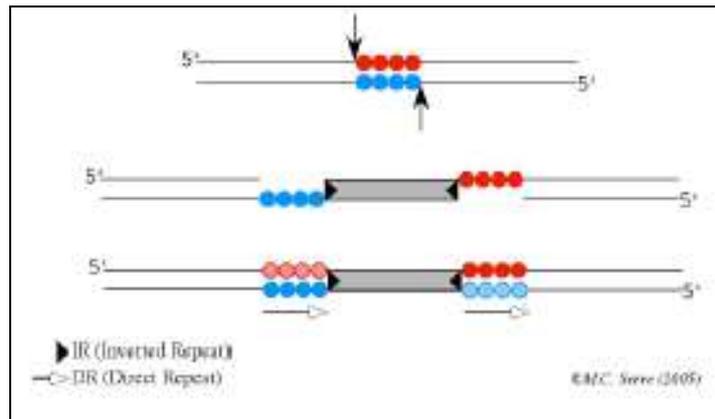
La transposition est déterminée par une ou plusieurs protéines spécifiques de chaque élément transposable. Tous les ET comprennent dans le segment d'ADN mobilisé un gène codant une protéine essentielle, **la transposase**. Les ET sont donc des éléments autonomes parasites des génomes dans lesquels ils se propagent. La transposase va reconnaître deux catégories distinctes de sites, d'une part les extrémités de l'élément et d'autre part un site cible (non spécifique) situé à un autre locus du génome. De plus, l'événement de recombinaison n'est pas réciproque, et est en général suivi d'une étape de synthèse d'ADN consécutive aux étapes de coupure et ligation essentielles pour l'échange des brins d'ADN.

- **Duplication de l'ADN au site d'insertion**

La signature quasi universelle d'un événement de transposition est la duplication de quelques paires de bases du site cible de l'insertion. Cette duplication est de longueur constante pour un ET donné, et varie de 2 à 14 pb.

Les deux copies de la séquence dupliquée se retrouvent en répétition directe encadrant l'élément intégré. La génération de répétitions directes (DR) est une conséquence du mécanisme d'intégration, et n'a aucun rôle à jouer dans des événements de transposition ultérieurs.

L'apparition des DR peut aisément s'expliquer par un modèle où la séquence cible de l'insertion va subir une coupure décalée, et où seul un brin de chaque extrémité de l'élément transposable serait répliqué aux extrémités sortantes. Les parties simple brin seront réparées par une néo synthèse d'ADN, générant les duplications aux bornes de l'élément nouvellement acquis.



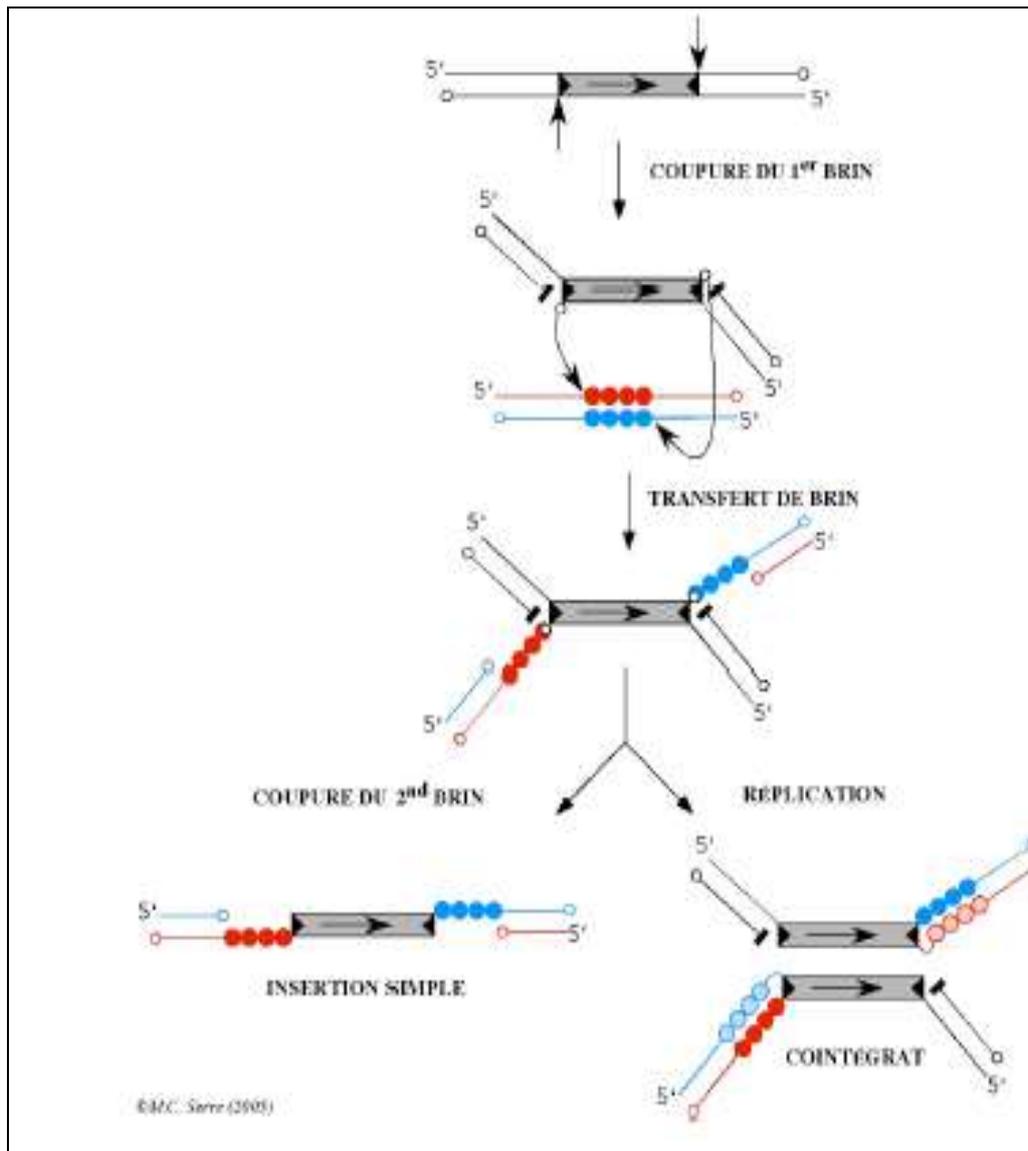
- **Transposition rélicative**

La présence de DRs aux bornes de l'ET suggère qu'un seul brin de chaque extrémité de l'ET serait transférée au cours de l'échange de brin avec la cible. Le devenir du second brin de l'ET est déterminant pour la suite de la réaction, et va définir le type de produit de transposition obtenu.

Si le brin des extrémités qui n'est pas transféré est laissé intact, on aboutit à une structure intermédiaire où la cible et le donneur sont réunis par l'élément. Cette structure appelée « **intermédiaire de Shapiro** » présente une structure similaire à une fourche de réplication à chaque extrémité du transposon. Sa résolution par réplication va conduire à la duplication du transposon, à la réparation du site cible et à la fusion des deux molécules (donneuse et accepteuse de l'ET). Le produit de cette réaction est nommé coïntégrat, et est caractéristique de la transposition de nombreux éléments bactériens. La résolution de l'intermédiaire de Shapiro est totalement dépendante de la machinerie de réplication de l'hôte.

- **Transposition non rélicative**

Un grand nombre d'éléments procaryotes ou eucaryotes ne se répliquent pas au cours de leur transposition. Le brin des extrémités qui n'est pas ligaturé à la cible est alors coupé à une étape précoce de la réaction. Le produit obtenu sera réparé aux bornes du transposon (ce qui va générer les DR), le résultat étant l'insertion «simple» de l'élément à un endroit quelconque du génome. Ce mode de transposition est fréquemment appelé « **couper-coller** ». (cut and paste). Dans tous les cas, cette forme est la preuve que les deux brins de chaque extrémité de l'ET ont été coupés **avant** le transfert dans la cible.



Ces éléments transposables constituent de véritables "**gènes sauteurs**", capables de se déplacer, par exemple, d'un plasmide sur le génome d'un bactériophage, puis sur le chromosome bactérien, puis à nouveau sur un plasmide, et ainsi de suite.

En raison de ces phénomènes, les éléments transposables, associés à des vecteurs tels que des plasmides sexuels ou des phages, constituent un élément essentiel dans l'évolution bactérienne. En particulier, ils contribuent de façon importante à la dissémination des gènes de **résistance aux antibiotiques**.

## II) Les Intégrons

Les gènes qui codent la résistance aux antibiotiques sont souvent localisés très près les uns des autres. L'analyse des séquences adjacentes à ces gènes de résistance montre en amont la présence d'un gène *intI* codant pour une intégrase suivi d'un site d'intégration *attI* unique et proche d'un site promoteur P. La séquence *attI* fonctionne comme un "hot spot" de recombinaison pour des séquences de DNA non homologues.

Ce n'est qu'au cours des années 1980, que des éléments génétiques susceptibles d'acquérir ou de perdre des gènes de résistance aux antibiotiques ont été décrits et désignés sous le nom d'**intégrons**.

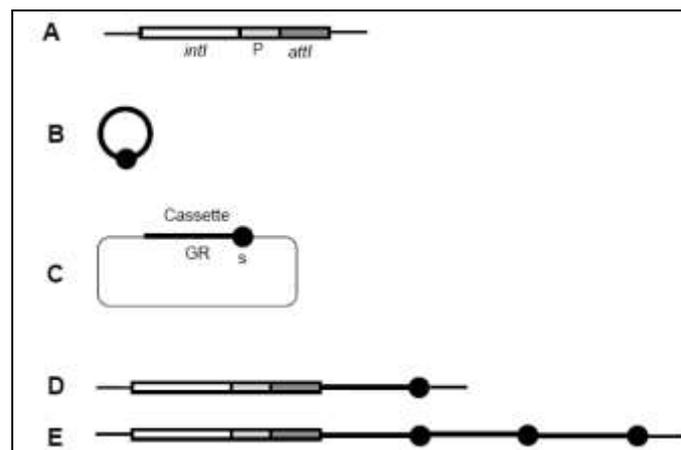
### 1) Définition

Les intégrons constituent un système de capture et d'expression de gènes sous forme de cassettes. Les **cassettes** sont des éléments mobiles capables d'être intégrés ou excisés par un mécanisme de recombinaison spécifique de site médié par une intégrase. Ces cassettes peuvent être présentes dans la cellule soit sous forme de DNA circulaire, soit intégrées dans un réplicon, comme un plasmide.

### 2) Structure des intégrons

Incapables d'autoréplication, les intégrons sont obligatoirement portés par un réplicon (plasmide ou chromosome). Ils peuvent aussi être véhiculés par un élément transposable. Les intégrons sont constitués d'une **région 5'** comprenant un gène *intI* qui code pour une **intégrase**, d'un site d'attachement *attI* et d'un promoteur (**P**). Le site *attI* permettra d'intégrer une ou plusieurs cassettes qui seront transcrites grâce à la présence du promoteur.

Les intégrons constituent donc un autre moyen efficace pour transmettre la résistance aux antibiotiques entre bactéries qui ne sont pas phylogénétiquement très proches les unes des autres.



**Figure:** Système des intégrons. A: Structure de base; B: Cassette sous forme circulaire; C: Cassette intégrée dans un plasmide (GR, gène de résistance; s, séquence "59 paires de bases"); D: Intégron avec une cassette; E: Intégron avec 3 cassettes.

Il existe **plusieurs classes d'intégrons** définies en fonction de la nature des gènes codant pour l'intégrase. Trois d'entre elles (classes 1, 2 et 3) ont été bien caractérisées et sont impliquées à

ce jour dans la dissémination de la résistance aux antibiotiques. Les sites *attI* des différentes classes d'intégrons n'ont pas de séquence commune, exceptée le motif GTTRRRY (R, purine ; Y, pyrimidine). Chez la majorité des intégrons de classe 1, la **région 3'** contient trois cadres de lecture ouverts. Le premier, *qacEDI*, est un dérivé tronqué du gène *qacE* codant pour la résistance aux ammoniums quaternaires. Le second cadre de lecture est le gène *sullI* qui code pour la résistance aux sulfamides. Le troisième cadre de lecture désigné *ORF5* ne code pour aucune fonction connue. Les intégrons diffèrent des transposons par plusieurs caractéristiques :

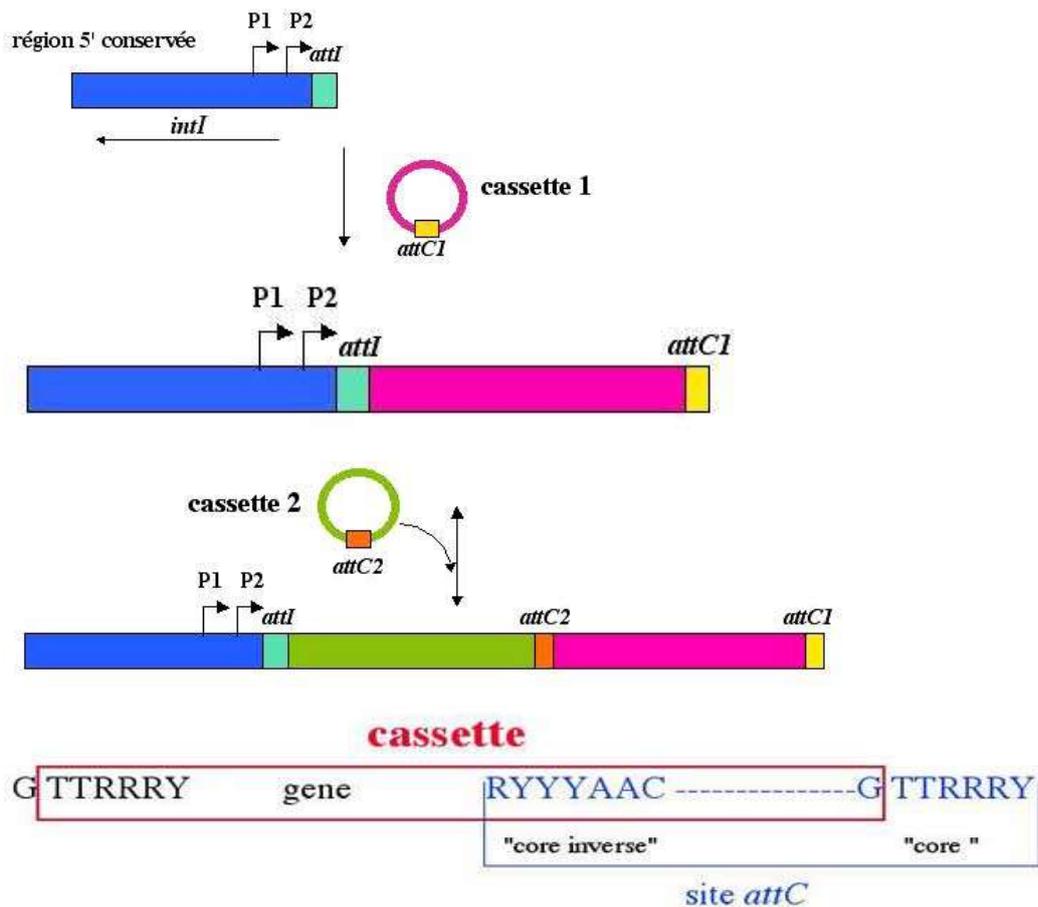
- les cassettes ne contiennent aucun gène codant pour une protéine catalysant leur mouvement, la recombinase étant présente sur la partie immobile de l'intégrons
- les cassettes ne sont pas flanquées à leurs extrémités de séquences inversées répétées.

### 3) Structure des cassettes

Les cassettes ont des tailles et des fonctions très variables mais possèdent une organisation commune. Une cassette est constituée d'un gène adjacent à un site spécifique de recombinaison *attC* reconnu par l'intégrase. Le site *attC* est constitué de séquences, relativement conservées, inversées répétées imparfaites dont la taille varie de 57 à 141 paires de bases. Deux séquences inversées répétées de 7 paires de bases sont constamment retrouvées aux deux extrémités de chaque site *attC* et désignées core et core inverse.

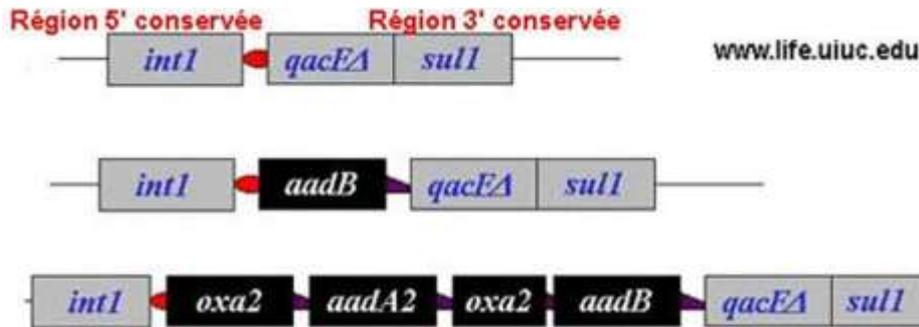
Le core de séquence consensus GTTRRRY est localisé à l'extrémité droite du site *attC* et le core inverse de séquence complémentaire RYYAAC à l'extrémité gauche.

Il existe une grande variété de sites *attC*. Certains, dont la séquence est très conservée, sont associés à des gènes de résistance très différents, alors que certains gènes apparentés sont associés à des sites *attC* hétérologues.



#### 4) Mouvement des cassettes

Certaines cassettes ont été retrouvées dans différentes classes d'intégrons. Toutes les cassettes peuvent apparemment être intégrées dans les trois classes d'intégrons. Cependant, les études sur le mouvement des cassettes ont principalement porté sur les intégrons de classe 1. Il a été démontré que les cassettes intégrées sous forme linéaire peuvent après excision, générer une forme libre circulaire. Le mouvement des cassettes se fait donc essentiellement par insertion-excision sous forme circulaire par un mécanisme de recombinaison entre deux sites spécifiques catalysé par l'intégrase. L'évènement de "crossing-over" se produit entre le G d'un site core GTTTRRY et le premier T d'un deuxième site core. Les intégrations de cassettes se font préférentiellement au site *attI*.



## 5) Expression des cassettes

Les cassettes sont toutes orientées dans la même direction et sont généralement dépourvues de promoteur. Les gènes sont co-transcrits comme au sein d'un opéron sous la dépendance d'une région promotrice localisée dans la région 5' conservée de l'intégron. L'expression des cassettes est tributaire de la position par rapport au promoteur, les gènes localisés dans des cassettes éloignées étant faiblement exprimés. La pression de sélection antibiotique peut favoriser des réarrangements de cassettes afin de permettre le positionnement d'une cassette faiblement exprimée à une localisation plus proche du promoteur. De rares cassettes possèdent leurs propres promoteurs tels que *cmlA* et *cmlA2* responsables de la résistance au chloramphénicol.

Chez les intégrons de classe 1, deux promoteurs P1 et P2 ont été caractérisés. P2 est situé 90 paires de bases en aval de P1. Il a été montré expérimentalement que ces deux promoteurs P1 et P2 étaient fonctionnels.

**Enseignante Responsable Mm R. GHARZOULI FERTOUL**