



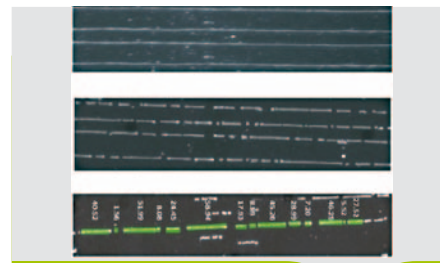
Méthodes

Les cartes optiques: principe et applications en microbiologie

M. Neto (1), G. Skorski (2), D. Thevenot (1), E. Loukiadis (1)

(1) VetAgroSup Campus vétérinaire de Lyon, Unité Calityss - Équipe EMSA-Lyon, LMAP - LNR STEC. 1 avenue Bourgelat. 69280 Marcy-l'Étoile, France

(2) Phylogène, 62 RN 113, 30620 Bernis, France



Les cartes optiques permettent d'obtenir une représentation graphique de la localisation des sites de restrictions dans le génome complet des organismes étudiés. Elles peuvent être obtenues de façon semi-automatisée par une méthode développée par la société OpGen. Les cartes optiques présentent de nombreux avantages pour l'aide et la finalisation du séquençage de génomes de procaryotes mais également d'eucaryotes. En épidémiologie, les cartes optiques sont un des seuls outils disponibles qui permettent d'étudier le génome complet; de plus, elles apportent une caractérisation très fine permettant de différencier les clones épidémiques des autres lignées clonales. Enfin, lors d'études en génomique comparative, les cartes optiques permettent de comparer aisément et rapidement les génomes de plusieurs bactéries et de mettre en évidence des insertions/délétions de gènes voire des réarrangements non encore décrits.

Introduction

Les cartes optiques permettent d'obtenir une représentation graphique de la localisation des sites de restriction dans le génome complet des organismes étudiés. Après avoir rappelé le principe d'obtention des cartes optiques, une « mini-revue » bibliographique des applications de ces cartes tant pour l'aide ou la vérification du séquençage, l'épidémiologie ou la génomique comparative sera présentée. Chacune de ces applications sera expliquée et illustrée notamment au travers d'exemples bibliographiques. Il est à noter que l'utilisation des cartes optiques pour l'étude de micro-organismes non cultivables ou l'identification d'organismes dans des échantillons complexes (échantillons cliniques...) ont aussi été rapportées (Mussmann *et al.* 2007; Skorski, communication personnelle).

Matériel et méthode

Principe

Le principe général des cartes optiques repose sur la réalisation d'une cartographie de la localisation de sites de restrictions sur le génome complet de l'organisme étudié (Figure 1-1).

OpGen est la seule société à réaliser une méthode semi-automatisée permettant, à partir d'un extrait d'ADN chromosomique, de réaliser l'ensemble des étapes nécessaires à l'obtention d'une carte optique (Figure 1-2).

L'extraction de l'ADN chromosomique des organismes étudiés est réalisée selon une méthode « douce » (en utilisant par exemple des billes magnétiques) afin d'obtenir de grands fragments d'ADN (Figure 1-1A). Ces brins sont ensuite immobilisés sur une lame spécifique (Figure 1-1B). Après vérification de son intégrité, de sa concentration (10^5 génomes bactériens en moyenne sont nécessaires à la réalisation d'une carte) et de l'absence de brins d'ADN < 150 kb, l'ADN fixé est ensuite digéré avec une enzyme de restriction adaptée à l'organisme et à l'objectif de l'étude (pour un même organisme plusieurs enzymes peuvent être utilisées indépendamment) (Figure 1-1C). L'enzyme de restriction utilisée est choisie à l'aide d'un logiciel (de type Enzyme Chooser -OpGen-) de façon à obtenir une taille moyenne de fragment comprise entre 6 et 12 kb et aucun fragment de restriction de plus de 80 kb. L'enzyme de restriction classiquement utilisée pour *Escherichia*

coli est *NcoI*. Après un marquage avec un fluorochrome (Figure 1-1D), les fragments d'ADN obtenus sont révélés par un microscope à fluorescence couplé à un logiciel d'analyse. La taille des fragments est mesurée (par comparaison avec un standard de taille) et l'ordre des fragments est pris en compte (Figure 1-1D).

Les fragments de taille inférieurs à 600 pb ne sont pas détectés. Les cartes de restriction unitaires sont ensuite assemblées en établissant des correspondances de profils de fragments pour produire la carte optique du génome entier de la bactérie étudiée (Figure 1-1E). La carte de restriction du génome complet alors obtenue ressemble à un code-barres (Figure 1-1F) où chaque barre correspond à un site de restriction. L'espace entre chaque barre correspond au nombre de paires de bases situées entre deux sites de restriction. Les cartes sont ensuite exploitées par des logiciels d'aide à l'interprétation de type MapSolver™.

Limites

Une des limites de cette technique est aujourd'hui sa disponibilité, car l'accès à cette technologie n'est possible de façon semi-automatisée qu'auprès de la société OpGen.

Par ailleurs, le coût de réalisation d'une carte est d'environ 380 € (prix communiqué par OpGen), celui de l'automate et des logiciels intégrés s'élève à environ 260 k€ et sa maintenance environ à 21 k€/an.

La technique des cartes optiques ne permet pas d'analyser les brins d'ADN de taille inférieure à 150 kb environ (tel que certains plasmides), car ces brins ne comportent pas suffisamment de sites de restriction pour créer un alignement fiable.

L'efficacité et les limites des cartes optiques notamment en ce qui concerne la détection des fragments de petites tailles ont été testées en comparant les cartes de restriction obtenues *in silico* de souches séquencées (appelées par la suite cartes *in silico*) et les cartes optiques réelles de ces mêmes souches (notamment pour *E. coli* O157:H7 (Kotewicz *et al.*, 2007) et *Yersinia pestis* (Zhou *et al.*, 2002)).

Par exemple, la comparaison des cartes optiques et des cartes *in silico* des deux souches séquencées d'*E. coli* O157:H7 EDL933 et Sakai a permis de montrer que, sur les 639 fragments *in silico* obtenus pour la souche Sakai, seuls



Méthodes

96 fragments ne sont pas détectés par la carte optique et que tous sont inférieurs à 2 kb (la majorité - 60/96 - sont d'ailleurs inférieurs à 600 pb car ces fragments sont trop petits pour être maintenus par la surface de la carte). Ces fragments non détectés représentent ainsi uniquement 20 309 bp sur un total de 5,5 Mb (soit environ 0,37 %) (Kotewicz *et al.*, 2007). Ces travaux montrent également que les 12 différences obtenues pour la souche EDL933 entre la carte *in silico* et la carte optique sont dues à des polymorphismes des sites de restrictions de BamHI ou à des défauts de coupures à six positions. Ces différences pourraient ainsi être attribuées soit aux limites technologiques des cartes (présence de fragments

de petites tailles non détectés) soit à des polymorphismes dans le génome de la souche étudiée. Néanmoins, les auteurs concluent que la méthodologie des cartes optiques reste un outil précis et utile pour différencier et étudier des souches très clonales.

Une aide pour le séquençage de génomes complets

L'assemblage des séquences représente une part importante des coûts et délais des séquençages. Les cartes optiques sont de plus en plus utilisées pour réduire ces délais et augmenter la

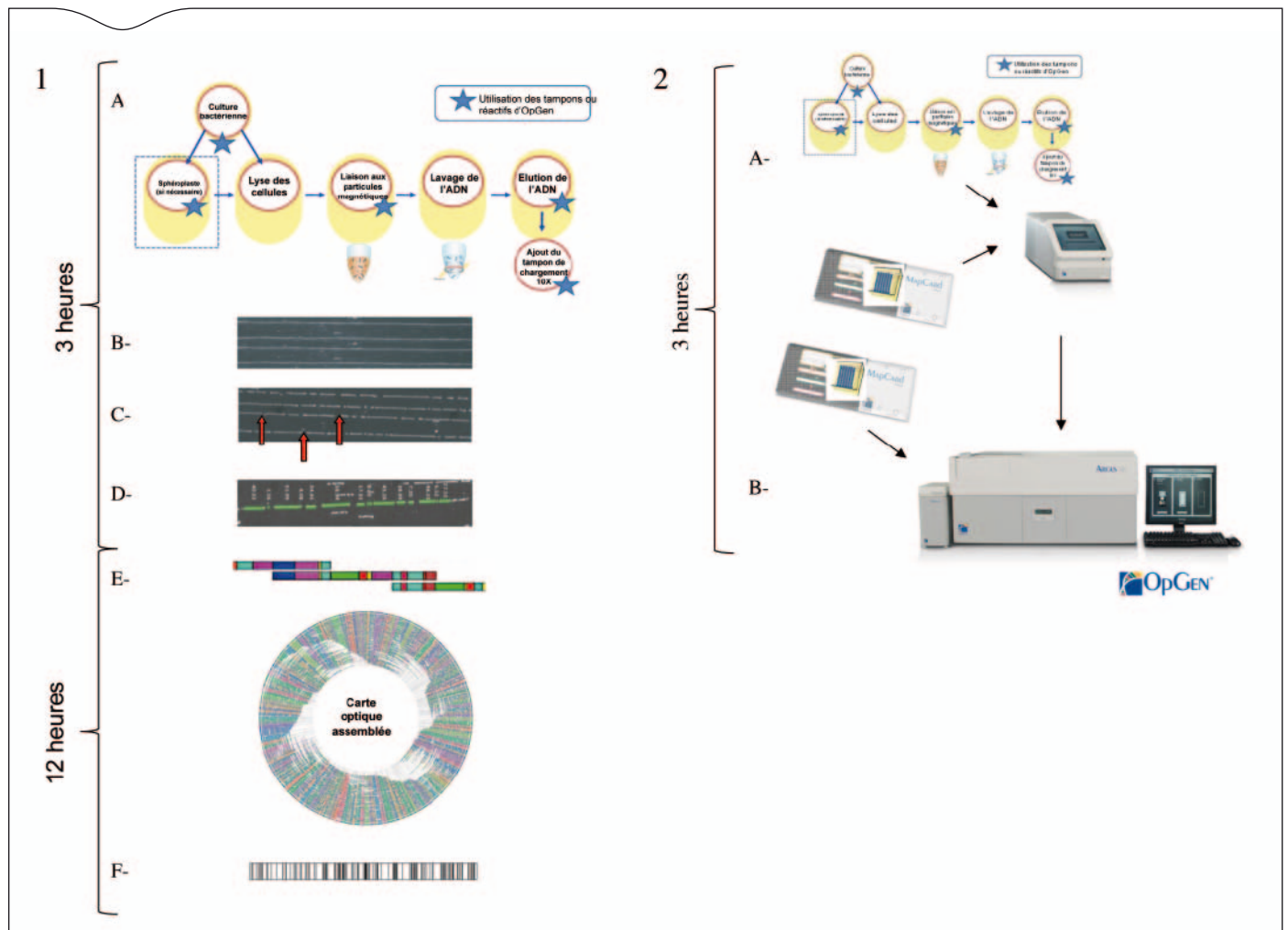


Figure 1. Méthode développée pour les micro-organismes par OpGen pour la réalisation d'une carte optique (avec l'autorisation d'OpGen)

1. Principe de la méthode :

A- L'ADN est extrait des cellules bactériennes, selon une méthode basée sur l'utilisation de billes magnétiques de façon à obtenir de grands fragments d'ADN. B- L'ADN est fixé sur un support spécifique. C- L'ADN est digéré avec une enzyme de restriction adaptée (les flèches rouges indiquent les sites de restriction) et marqué par un fluorochrome. D- L'intensité de fluorescence est mesurée pour déterminer la taille des fragments et leur ordre est enregistré. E- Les brins d'ADN qui se recouvrent sont assemblés pour produire une carte optique du génome complet. F- La carte optique est représentée ici, chaque ligne verticale représente la localisation d'un site de restriction et l'espace entre les barres représentent la taille en kilobases des fragments de restrictions obtenus.

2. Méthode automatisée développée pour les microorganismes par la société OpGen :

A- L'ADN obtenu est déposé sur le support MapCard™, quantifié et digéré dans l'automate « Argus™ Mapping System » permettant la création d'une carte. B- La carte ainsi créée est lue et analysée de façon automatique par le système « Argus™ Optical Mapping System ».



Méthodes

fiabilité de l'assemblage de séquences de génomes complets d'organismes procaryotes mais aussi eucaryotes. Le Tableau 1 reprend ainsi des exemples tirés de la bibliographie.

La réalisation de cartes optiques d'un micro-organisme en cours de séquençage permet i) de connaître la taille du génome; ii) de servir de base à l'assemblage des séquences; iii) de confirmer, en comparant la carte optique réalisée sur l'ADN de l'organisme et la carte réalisée *in silico*, le bon alignement et/ou de mettre en évidence des gaps sur la séquence finale (la Figure 2 illustre de façon schématique l'aide apportée par les cartes optiques pour la résolution et la vérification des assemblages d'un séquençage de génome complet (modifiée de Aston *et al.*, (1999) et Neely *et al.* (2010)). La carte *in silico* de la séquence contigüe obtenue à partir du séquençage est comparée à la carte optique de la souche bactérienne étudiée. Cette comparaison permet de confirmer l'assemblage et le sens des séquences contigües et de mettre en évidence les « gaps » et d'aider à leur résolution); iv) de vérifier le séquençage des génomes possédants de nombreuses insertions de prophages, ou des régions présentant des séquences répétées qui sont souvent sources de mésappariements avec les outils traditionnels (Jing *et al.* 1999; Lim *et al.* 2001).

La réalisation de deux cartes optiques avec deux enzymes de restriction, fournissant les résultats de deux cribles et mettant en évidence la localisation et la répartition de deux types de sites de restrictions, peut également apporter des informations complémentaires et aider à l'assemblage des séquences (cette démarche a d'ailleurs été utilisée pour le séquençage de *Yersinia pestis* par Zhou *et al.* (2002).

De plus, les cartes optiques nécessitant une faible quantité

d'ADN (10^5 génomes bactériens en moyenne), elles peuvent être utilisées pour les organismes dont la culture et le clonage sont délicats tels que *Plasmodium falciparum* (Jing *et al.* 1999). En outre, par rapport aux autres outils d'aide à l'assemblage des séquences tels que les cartes de restriction ou les hybridations, les cartes optiques présentent l'avantage de ne pas nécessiter d'étapes pouvant induire d'artefacts (tels que les biais liés aux clonages parfois nécessaires avant le séquençage) et de pouvoir être réalisées en parallèle et indépendamment du séquençage.

Les travaux de Giongo *et al.* (2010) relatifs à *Gluconacetobacter diazotrophicus* illustrent d'ailleurs particulièrement ce point. En effet, deux séquençages indépendants d'une même souche PAI5 présentaient des divergences et les cartes optiques réalisées ont permis de conforter une des séquences de la souche PAI5.

Ainsi, les cartes optiques semblent être un outil fidèle et particulièrement intéressant pour l'aide ou la vérification des assemblages d'un génome séquencé (Giongo *et al.*, 2010; Zhou *et al.*, 2004b) notamment car elles font appel à un principe et des méthodes différents, indépendants et complémentaires au séquençage.

Un outil de typage et de comparaison de souches microbiennes utilisable en épidémiologie

De nombreux outils sont disponibles à l'heure actuelle pour étudier les relations clonales de souches isolées lors d'épidémies ou de poly-infections (électrophorèse en champ pulsé (PFGE); Multilocus sequence typing (MLST), Multiple-

Tableau 1. Exemples de travaux utilisant ou présentant les cartes optiques comme une aide au séquençage de génomes complets

Organismes concernés	Utilisations	Références
Grands génomes et eucaryotes		
<i>Aspergillus oryzae</i>	Vérification de la séquence	(Machida <i>et al.</i> , 2005)
<i>Aspergillus fumigatus</i>	Aide à l'assemblage	(Nierman <i>et al.</i> , 2005)
<i>Candida albicans</i>	Vérification de la séquence	(van het Hoog <i>et al.</i> , 2007)
Homme (Human BAC)	Aide à l'assemblage	(Cai <i>et al.</i> , 1998)
<i>Leishmania major</i> Friedlin	Aide à la résolution de gaps et erreurs de contigs	(Zhou <i>et al.</i> , 2004b)
<i>Plasmodium falciparum</i> (parasite)	Organisme difficilement cultivable; besoin de peu d'ADN pour la réalisation de la carte optique; aide à l'assemblage	(Jing <i>et al.</i> , 1999)
	Aide à l'assemblage	(Lai <i>et al.</i> , 1999)
Procaryotes		
<i>Arthrospira platensis</i>	Aide à l'assemblage	(Fujisawa <i>et al.</i> , 2010)
Large Sulfur Bacteria (<i>Beggiatoa</i>)	Aide à l'analyse de bactérie non cultivable	(Mussmann <i>et al.</i> , 2007)
<i>Escherichia coli</i> O157:H7	Aide à l'analyse des contigs par des informations sur les réarrangements	(Lim <i>et al.</i> , 2001)
<i>Gluconacetobacter diazotrophicus</i> PAI 5	Comparaison du séquençage de la même souche par deux équipes. Montre ainsi les erreurs possibles d'assemblages et/ou réarrangements rapides. Intérêts des cartes optiques pour la proposition de nouvelles séquences.	(Giongo <i>et al.</i> , 2010)
<i>Xenorhabdus nematophila</i> et <i>Xenorhabdus bovienii</i>	Aide à l'assemblage	(Latreille <i>et al.</i> , 2007)
<i>Yersinia pestis</i>	Utilisation de deux enzymes différentes: une pour aide à l'assemblage, l'autre pour la vérification de la séquence. Mise en évidence d'erreurs d'assemblages et aide au bon assemblage par la carte optique.	(Zhou <i>et al.</i> , 2002)

Méthodes

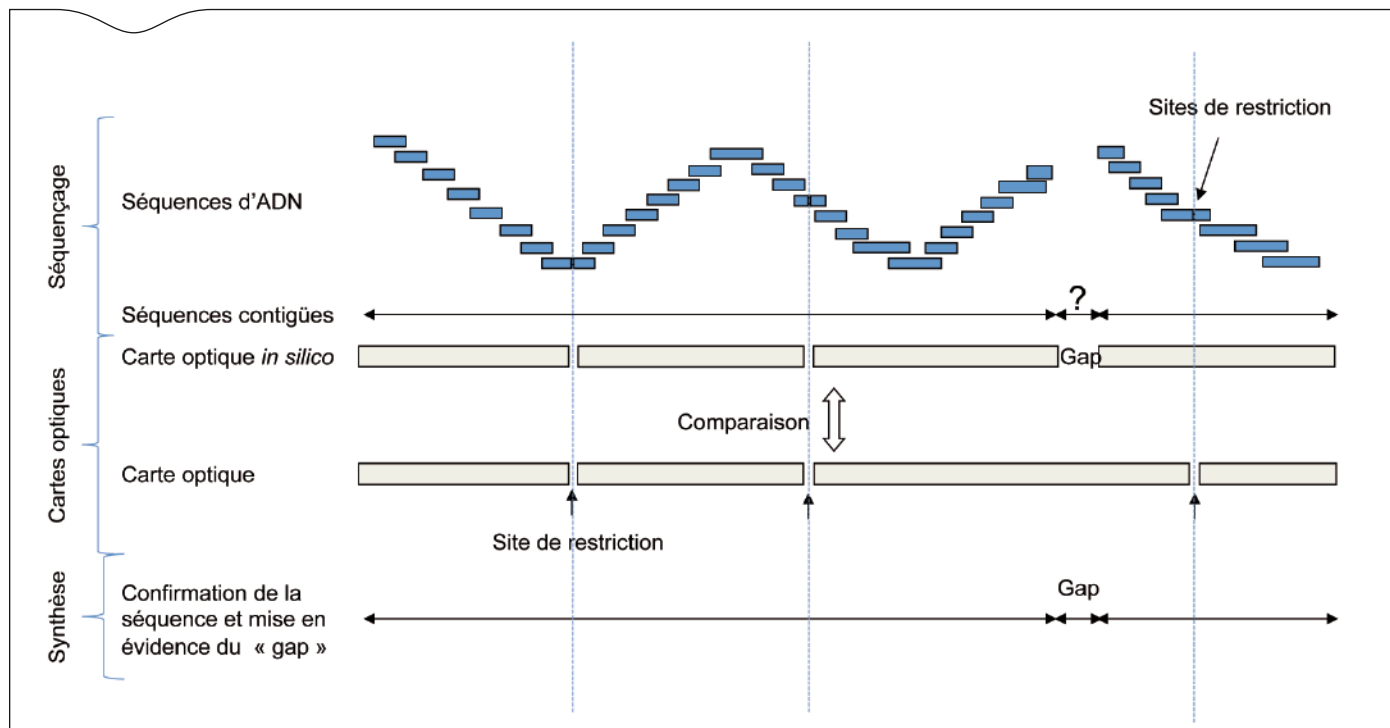


Figure 2. Représentation schématique de l'aide apportée par les cartes optiques pour la résolution et la vérification des assemblages d'un séquençage de génome complet (modifiée de Aston *et al.*, (1999) et Neely *et al.* (2010)).

La carte in silico de la séquence contigüe obtenue à partir du séquençage est comparée à la carte optique de la souche bactérienne étudiée. Cette comparaison permet i) de confirmer l'assemblage et le sens des séquences contigües; ii) de mettre en évidence les « gaps » et d'aider à leur résolution. Les cartes optiques peuvent également servir d'échafaudage à l'assemblage des séquences pour obtenir la séquence contigüe.

locus variable-number tandem-repeats analysis (MLVA), repetitive-sequence-based PCR (rep-PCR)...). Parmi eux, seuls trois permettent d'étudier le génome complet: la PFGE, la rep-PCR et les cartes optiques.

La rep-PCR semble présenter le même pouvoir discriminant que la PFGE pour certains organismes (tels que *Salmonella enterica* (Foley *et al.* 2006) et *Clostridium difficile* (Northey *et al.* 2005) mais être moins discriminante pour d'autres organismes tels que *E. coli* O157:H7 (Foley *et al.* 2004). Les outils disponibles et applicables aux *E. coli* producteurs de Shiga-toxines (STEC) ont d'ailleurs été comparés récemment dans une revue (Karama and Gyles, 2010). Les auteurs indiquent aussi que les cartes optiques et les PFGE présentent les mêmes performances en termes de pouvoir discriminant, de facilité de réalisation et de standardisation mais que toutes deux présentent des difficultés d'interprétation et une analyse des résultats laborieuse. Toutefois, en ce qui concerne le délai et la complexité de l'analyse des données obtenues par cartes optiques (par comparaison à la PFGE), il convient de prendre en compte l'objectif poursuivi. En effet, si l'on souhaite étudier les distances clonales entre souches, l'analyse des données obtenues par les cartes optiques n'est ni plus longue ni plus complexe que celle obtenues par PFGE. De plus, si l'on étudie les caractéristiques de clones épidémiques ou de groupes de souches (ce qui est difficilement réalisable avec la PFGE), les cartes optiques semblent être plus discriminantes, en raison du nombre plus important de fragments obtenus et analysés par carte optique comparativement à la PFGE. En effet, Kotewicz

et al. (2007) montrent que les cartes optiques présentent une résolution et des détails suffisants pour différencier des souches *E. coli* O157:H7 très proches. En effet, la comparaison du nombre de fragments obtenus par une PFGE avec l'enzyme *Xba*I classiquement utilisée pour *E. coli* O157:H7 (43 fragments) et avec celui obtenu en carte optique avec l'enzyme *Bam*HI (645 fragments) suggèrent que les cartes optiques seraient plus informatives.

Dans le cas de la souche séquençée 11368 d'*E. coli* O26:H11 (Figure 3), le nombre de sites de restriction obtenus avec l'enzyme *Nco*I, utilisée pour la réalisation des cartes optiques (725 fragments, Figure 3A) est supérieur à celui obtenu avec l'enzyme *Xba*I utilisée classiquement en PFGE (36 fragments, Figure 3B), le choix de l'enzyme de restriction pour cette dernière étant en effet limité par la visualisation de fragments distincts. En considérant le nombre de fragments obtenus dans cet exemple précis, les résultats des cartes optiques donnent un niveau de détail 20 fois supérieur à ceux obtenus par la PFGE.

Au laboratoire, huit souches d'*E. coli* O26:H11 d'origine alimentaire et la souche O26 séquençée 11368 d'origine humaine ont été typées et comparées par PFGE (Figure 3D) et par cartes optiques (Figure 3E). Les deux principaux clusters obtenus par ces deux techniques sont similaires et permettent de conclure que les souches 108.1.1, 60.1 et 64.1 appartiennent au même clone et forment un cluster distinct des autres souches d'*E. coli* O26:H11 étudiées. Cependant, les relations clonales entre les souches du cluster 1 (Figure 3) sont

Méthodes

différentes selon les deux techniques ce qui pourrait illustrer le plus grand pouvoir discriminant des cartes optiques par rapport à la PFGE.

Enfin, comparativement aux résultats obtenus par PFGE, les cartes optiques présentent l'avantage d'ordonner les fragments selon le génome et non selon la taille du fragment (Figures 3B et 3C) et d'indiquer la localisation des fragments de restriction sur le génome, ce qui peut se révéler particulièrement intéressant pour la caractérisation de clones épidémiques (Kotewicz *et al.*, 2008) ou de groupes de souches (Schwan *et al.*, 2010).

En effet, en réalisant les cartes optiques des isolats d'*E. coli*

O157:H7 provenant d'épinards contaminés et à l'origine d'une épidémie alimentaire aux USA, Kotewicz *et al.* (2008) ont mis en évidence 14 marqueurs chromosomiques associés à un clone épidémique notamment grâce à la comparaison des cartes *in silico* des génomes d'EDL933 et Sakai séquencés. En associant des analyses de séquençages et les cartes optiques, les auteurs ont pu montrer que le clone épidémique incriminé ne possédait pas le gène *stx1* bien que possédant un prophage similaire au prophage V de EDL933 portant classiquement ce gène. De plus, l'analyse des cartes optiques obtenues couplée à un séquençage des marqueurs chromosomiques potentiellement intéressants a permis de localiser dans le génome les gènes *stx2* et *stx2c*. Ces résultats suggèrent que la virulence particulière de certaines souches d'*E. coli* O157:H7 pourrait ainsi dépendre de l'emplacement ou des réarrangements des séquences spécifiques des phages.

Une autre étude portant sur les *E. coli* uropathogéniques (UPEC) (Schwan *et al.*, 2010) a montré, grâce à l'utilisation des cartes optiques, que deux groupes d'UPEC peuvent être différenciés. Ces deux groupes correspondent à des profils de présence/absence de facteurs de virulence et aux profils d'antibiorésistance observés chez les souches. Les cartes optiques ont été présentées par les auteurs comme un outil possible de détection rapide de souches résistantes aux antibiotiques lors d'épidémies.

D'autres exemples d'études épidémiologiques (recensées entre 2006 et 2010) utilisant les cartes optiques sont cités dans le Tableau 2 et soulignent toutes l'intérêt de leur utilisation pour l'étude précise de groupes de souches et de clones épidémiques. Par ailleurs, les cartes optiques permettant également de reconstruire des génomes en alignant des brins d'ADN (avec une limite de détection d'environ 105 germes), elles peuvent être utilisées pour l'étude d'organismes non cultivables (Mussmann *et al.*, 2007) ainsi que pour l'étude d'échantillons poly-microbiens complexes (échantillons cliniques...) afin d'identifier de nouvelles souches voire d'étudier la coopération entre souches (Skorski, communication personnelle).

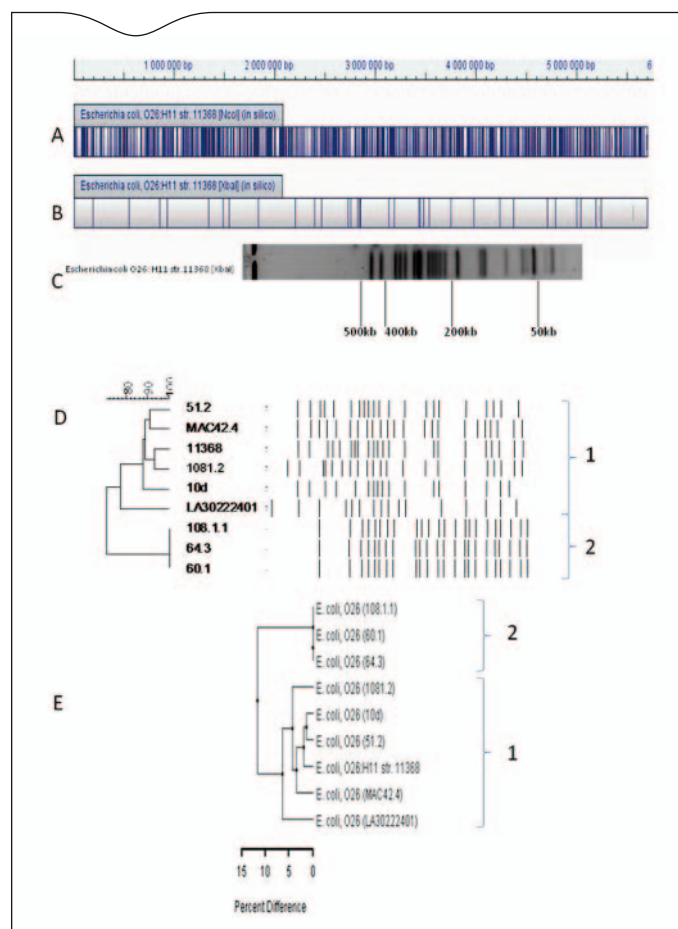


Figure 3. Cartes optiques, PFGE et clusters de similarité de souches d'*E. coli* O26:H11.

A : Carte optique *in silico* de la souche séquencée 11368 d'*E. coli* O26:H11 isolée d'un patient atteint de diarrhée (Ogura *et al.*, 2009) avec l'enzyme NcoI (725 fragments). B : Carte optique *in silico* de la souche 11368 de *E. coli* O26:H11 avec l'enzyme XbaI (36 fragments). C : PFGE obtenue avec l'enzyme XbaI de la souche 11368 d'*E. coli* O26:H11. D : Cluster de similarité obtenu par PFGE des 9 souches d'*E. coli* O26:H11 d'origine humaine ou alimentaire. Restriction avec l'enzyme XbaI. Dendrogramme construit à l'aide de GelComparII (Applied-Maths) en utilisant le coefficient de Dice avec une tolérance de 1,5 % sur la position des bandes et UPGMA (unweighted pair group method using arithmetic averages), l'échelle représente le pourcentage de similarité. E : Cluster de similarité obtenu de l'analyse de cartes optiques de neuf souches d'*E. coli* O26:H11 d'origine humaine ou alimentaire. Dendrogramme construit à partir des cartes optiques (réalisées avec l'enzyme NcoI) en utilisant UPGMA, l'échelle représente le pourcentage de différence.

Tableau 2. Exemples de travaux utilisant ou présentant les cartes optiques pour des études d'épidémiologie

Organismes concernés	Utilisations	Références
<i>E. coli</i> enterotoxigéniques (ETEC)	Comparaison de souches <i>E. coli</i> K12 et ETEC H10407. Souches proches.	(Chen <i>et al.</i> , 2006)
<i>E. coli</i> uropathogéniques (UPEC)	Comparaison entre les UPEC, <i>Shigella</i> , EHEC et souche K12. Les UPEC forment un groupe indépendant. Au sein des UPEC, 2 groupes corrélés à la présence/absence de facteurs de virulence et de résistances aux antibiotiques.	(Schwan <i>et al.</i> , 2010)
<i>E. coli</i> O157:H7	Comparaison de souches entre elles et étude en génomique comparative (cf. <i>infra</i>).	(Kotewicz <i>et al.</i> , 2007)
<i>E. coli</i> O157:H7	Étude d'un clone épidémique. Mise en évidence de 14 marqueurs chromosomiques spécifiques (cf. <i>infra</i>).	(Kotewicz <i>et al.</i> , 2008)



Méthodes

Les outils de typage disponibles et permettant l'étude des génomes complets sont limités (PFGE, rep-PCR, cartes optiques) et présentent chacun des avantages et des inconvénients en terme de faisabilité, de disponibilité, de pouvoir discriminant et de coûts. Il s'agit ainsi d'adapter l'outil utilisé à l'objectif poursuivi notamment quant à la précision des résultats souhaités, mais aussi aux délais souhaités d'obtention des résultats et au budget disponible. En effet, si les cartes optiques sont moins disponibles (technologie uniquement accessible par le biais de la société OpGen) et plus coûteuses que des outils comme les PFGE et la rep-PCR, elles sont plus informatives car plus discriminantes que la PFGE (20 fois plus de fragments obtenus par exemple pour la souche 11368 d'*E. coli* O26:H11) et surtout elles permettent d'accéder au positionnement et à l'ordre des fragments de restriction dans le génome. Ces dernières caractéristiques en font un outil très précis et utile notamment pour l'analyse fine de clones épidémiques.

Un outil pour la génomique comparative du génome complet: détection d'insertions, délétions, inversions et translocations

Les outils de typages, de PCR et de séquençages de génomes bactériens complets ont permis de mettre en évidence la grande plasticité génomique (réarrangements, délétions, insertions) des micro-organismes à l'origine de leur diversité et de leur évolution. Or, les cartes optiques peuvent présenter des avantages par rapport à ces outils classiquement utilisés en génomique comparative :

- les puces à ADN et les hybridations permettent de cribler un grand nombre de gènes et ce, rapidement et sur un grand nombre de souches; mais, contrairement aux cartes optiques, permettent uniquement de détecter la présence ou l'absence de gènes et non d'étudier des gènes inconnus ou d'éventuels insertions et réarrangements génétiques;
- le typage par macro-restriction (type PFGE) met en évidence les différences entre les génomes mais, contrairement aux cartes optiques, ne permet pas de relier directement ces différences à des gènes ou à une position dans le génome;

- le séquençage, par une stratégie de « shotgun » ou haut débit, présente l'avantage d'être plus complet car il fournit plus d'informations (obtention de la séquence génétique complète) mais, à ce jour, les cartes optiques restent l'outil le plus adapté pour l'étude des espèces pour lesquelles peu de séquences sont disponibles.

Kotewicz *et al.* (2007) et Zhou *et al.* (2004a) ont montré la rapidité et la facilité de l'analyse des cartes optiques, notamment en raison de leur aspect visuel et graphique. De plus, la correspondance entre les zones d'intérêts (repérées lors de l'analyse des cartes) avec l'annotation des gènes est possible par la comparaison avec une carte *in silico* d'une souche séquencée (Figure 4). Néanmoins, si les cartes optiques permettent de mettre en évidence de nouveaux réarrangements et de nouveaux loci génétiques, le séquençage de la zone présentant ces nouveaux loci ou réarrangements reste nécessaire puisqu'il permet de confirmer les observations effectuées avec la carte optique (Shukla *et al.*, 2009).

De plus, il est à noter que le niveau de résolution des cartes optiques ne permet pas de mettre en évidence les mutations ponctuelles ou de faible ampleur.

Plusieurs études illustrant l'apport des cartes optiques en génomique comparative peuvent ainsi être citées.

La comparaison des cartes optiques des souches séquencées EDL933 et Sakai d'*E. coli* O157:H7 montrent seulement deux différences significatives en ce qui concerne la taille de certains fragments et l'existence d'une inversion de 430 kb (soit 7,8 % du génome) dans une région flanquée des prophages O et P chez EDL933 (Kotewicz *et al.*, 2007).

Les cartes optiques des souches EDL933 et Sakai ont également été comparées à neuf autres souches d'*E. coli* O157:H7 (Kotewicz *et al.*, 2007). Au total, 91 différences ont été mises en évidence; toutes regroupées dans 28 loci dont la majorité correspond à des prophages, à des prophages cryptiques ou probablement à de nouveaux sites d'insertion de phages. Cette étude a aussi permis de montrer que les deux sites les plus variables étaient des sites occupés par les prophages portant les gènes *stx1* et *stx2*.

La comparaison des cartes optiques de quatre sérovars de *Salmonella enterica* subsp *enterica* (Saunders *et al.* 2010) a

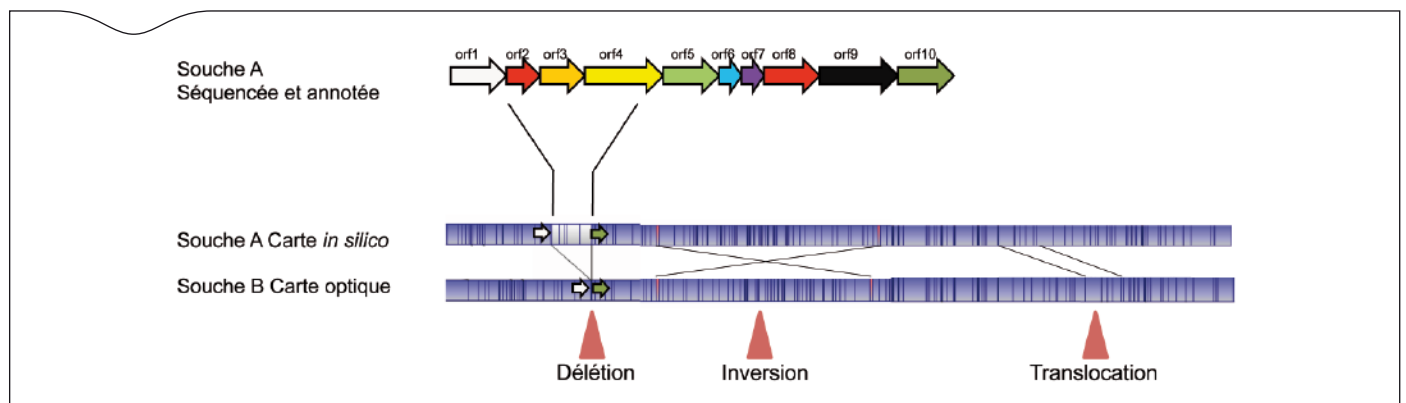


Figure 4. Représentation schématique de l'apport des cartes optiques en génomique comparative (comparaison d'une souche séquencée et d'une souche d'intérêt).

Les barres verticales dans les cartes optiques représentent les sites de restriction, les flèches représentent les ORFs. La carte optique de la souche A est réalisée *in silico* à partir de la séquence nucléotidique disponible dans les bases de données. L'utilisation combinée des cartes optiques et des séquences permet de mettre en évidence les différences entre les génomes telles que les insertions, translocations, inversions et délétions (représentées par des triangles).



Méthodes

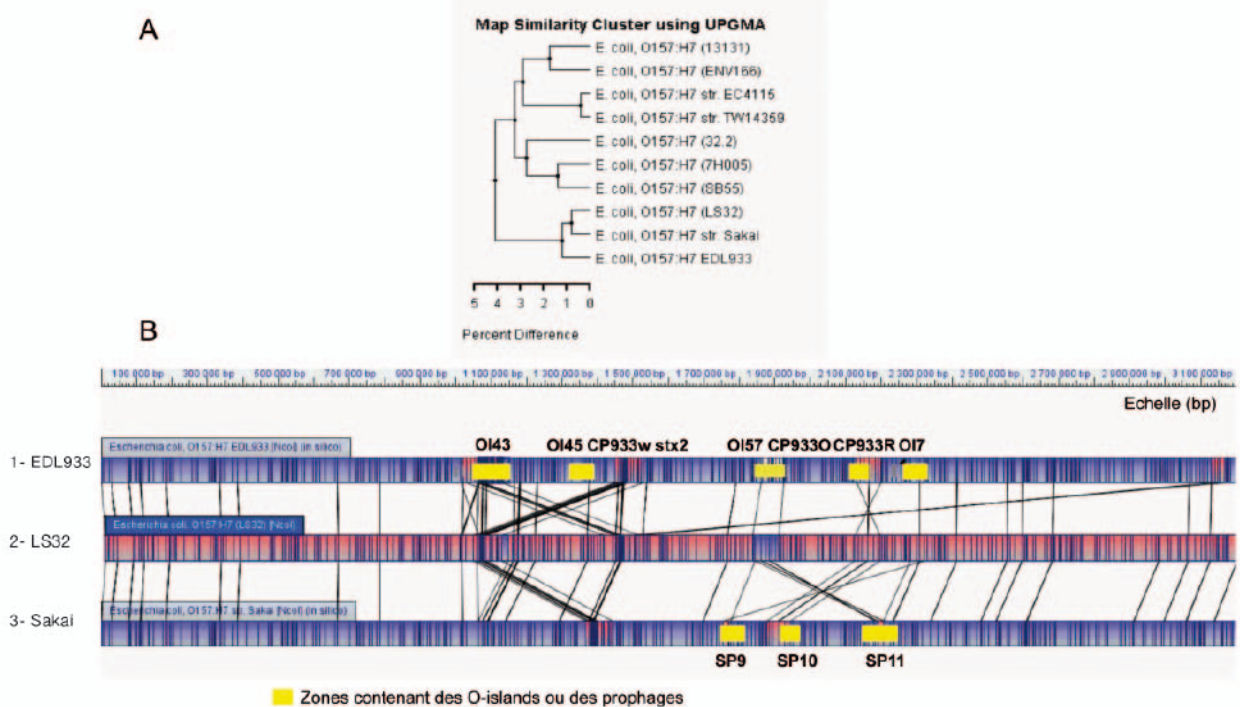


Figure 5. Comparaison de souches d'*E. coli* O157:H7.

A: Cluster de similarité obtenu à partir de l'analyse de cartes optiques de 10 souches d'*E. coli* O157:H7 d'origine humaine ou alimentaire réalisées avec l'enzyme NcoI. Dendrogramme construit à partir des cartes optiques (réalisées avec l'enzyme NcoI) en utilisant la méthode UPGMA. L'échelle représente le pourcentage de différence. Les cartes optiques des souches Sakai, EDL933, TW14359 et EC4115 ont été réalisées in silico à partir des séquences Genbank: NC_002695, NC_002655, NC_013008, NC_011353.

B: Exemple de la région la plus variable repérée par la comparaison de 3 souches: Sakai, EDL933 et LS32. Les cartes optiques sont réalisées avec l'enzyme NcoI: 1) carte in silico de la souche séquencée EDL933; 2) carte optique de la souche LS32; 3) carte in silico de la souche séquencée Sakai. Les souches EDL933 et LS32 sont divergentes au sein de la zone correspondante à O-Island 57 (CP933O) dans la souche séquencée EDL933. La zone est différente dans la zone codant pour un prophage criptique CP933R de EDL933 présente une inversion par rapport à la souche LS32. Une insertion est présente dans la souche LS32 par rapport à la souche Sakai, cette zone est annotée chez EDL933 comme correspondant à l'O-Island 43. L'inversion montrée par Kotewicz et al. (2007) entre les O-islands 57 et 71 entre les souches EDL933 et Sakai, n'est présente ici que chez la souche Sakai. La souche LS32 présente la même configuration que la souche EDL933. Cette zone comporte les OI57 (CP933O), CP933R, OI68, OI69, OI70 et OI71 (Perna et al., 2001). Cette zone code pour des prophages (SP (sakai prophage 9, SP10, SP11, SP12 (Hayashi et al., 2001)).

permis de mettre en évidence des variations dans le génome de *S. typhimurium*, associées à la présence de prophages et corrélées aux profils d'antibiorésistance observés, ce qui a permis d'envisager de nouvelles pistes de recherche.

Une étude en cours dans notre laboratoire (Neto et al. en préparation) vise à comparer des souches STEC O157:H7 d'origine alimentaire ou humaine par carte optique en collaboration avec la société OpGen. La comparaison des cartes optiques de la souche d'origine humaine LS32 (collection LMAP-LNR STEC) et de deux souches humaines séquencées (EDL933 et Sakai) (Figure 5A) a montré que ces souches étaient très proches en terme de structure génomique et que les divergences se situaient principalement dans une zone riche en « O-islands » et prophages (Figure 5B). Dans le cas d'une souche non séquencée, les cartes optiques permettent d'orienter rapidement la recherche par PCR de la position des prophages dans le génome. Ces résultats illustrent la contribution importante des prophages dans la plasticité

génomique des *E. coli* O157:H7. L'augmentation du nombre et de la diversité des micro-organismes étudiés par carte optique (en particulier sur un éventail plus large d'entérobactéries) pourrait permettre de détecter de nouvelles zones d'instabilité et des éléments génétiques mobiles impliqués dans l'émergence des *E. coli* pathogènes.

Le Tableau 3 présente une synthèse des autres études de génomique comparative conduites en ayant eu recours aux cartes optiques.

Les cartes optiques, en raison de leurs avantages techniques (pas d'information préalable sur le génome nécessaire, accès à l'information relative à la position et l'ordre des fragments de restriction sur le génome) permettent d'étudier des espèces pour lesquelles peu de données sont disponibles, de mettre en évidence la présence de nouveaux gènes et de localiser de nouvelles insertions et des réarrangements génétiques non encore décrits.



Méthodes

Tableau 3. Exemples de travaux utilisant ou présentant les cartes optiques pour des études en génomique comparative

Organismes concernés	Utilisations	Références
<i>Bifidobacterium animalis</i>	Comparaison de souches. Peu de différences entre les souches.	(Barrangou <i>et al.</i> , 2009)
<i>Enterococcus faecium</i>	Comparaison de souches et mise en évidence de différences entre clones épidémiques.	(Johnson <i>et al.</i> , 2010)
<i>E. coli</i> enterotoxigéniques (ETEC)	Comparaison de souches. Montre de fortes similitudes entre <i>E. coli</i> K12 et ETEC H10407	(Chen <i>et al.</i> , 2006)
<i>E. coli</i> O157:H7	Comparaison de souches et mise en évidence des régions variables, des insertions, délétions, inversions d'ADN entre les souches.	(Kotewicz <i>et al.</i> , 2007)
<i>E. coli</i> O157:H7	Mise en évidence de 14 marqueurs chromosomiques spécifiques d'un clone épidémique.	(Kotewicz <i>et al.</i> , 2008)
<i>Salmonella enterica</i> subsp <i>enterica</i>	Comparaison de souches. Montre que les délétions de certaines cassettes sont à l'origine de différents profils d'antibiorésistance. Mise en évidence de réarrangements génétiques non encore décrits.	(Saunders <i>et al.</i> , 2010)
<i>Shigella flexneri</i>	Comparaison de souches. Mise en évidence de nouveaux sites portant des séquences d'insertions et des insertions de gènes de phages.	(Zhou <i>et al.</i> , 2004a)

Tableau 4. Avantages et inconvénients de l'utilisation des cartes optiques dans les domaines présentés ici

Applications des cartes optiques	Avantages	Inconvénients
Aide et/ou vérification du séquençage de génomes	Indépendant du séquençage (méthode pouvant être mise en œuvre en parallèle) Ne nécessite pas de clonage ou d'autres techniques pouvant créer des artefacts Rapide et fiable	Disponibilité réduite* Pas d'information sur les plasmides
Epidémiologie	Analyse basée sur le génome complet Fort pouvoir discriminant (200 marqueurs par Mbases) Facilité de réalisation Facilité de standardisation Information relative à l'ordre des fragments dans le génome et comparaison possible avec des souches séquencées Accès à l'annotation des souches séquencées, et annotation possible des cartes optiques Fonction clusters intégrée	Peu de facilité d'interprétation et une analyse des résultats laborieux Disponibilité réduite* Coût** Pas d'information sur les plasmides
Génomique comparative	Rapidité et la facilité de l'analyse avec le logiciel Aspect visuel et graphique Comparaison aisée de nombreuses souches Possibilité de montrer de nouveaux réarrangements non décrits Information relative à l'ordre des fragments dans le génome et comparaison possible avec des souches séquencées Possibilité d'adaptation à d'autres techniques telles que les « DNA probings » (ou sondes marquées) Accès à l'annotation des souches séquencées, et annotation possible des cartes optiques Fonction clusters intégrée	Disponibilité réduite* Coût** Pas d'information sur les plasmides de taille inférieurs à 150 Kb

* La technologie est disponible chez OpGen.

** Le coût d'une carte est d'environ 380 € pour les réactifs. Le coût de l'automate et des logiciels intégrés à la machine s'élève à d'environ 260 k€ avec une maintenance annuelle de 21 k€.



Méthodes

Conclusion

Les cartes optiques peuvent être utilisées pour de nombreuses applications avec des avantages indéniables mais aussi des inconvénients liés au principe même de la méthodologie employée, à sa disponibilité (technologie accessible seulement par le biais de la société OpGen) et à son coût. Ces aspects, pour chaque application, sont synthétisés dans le Tableau 4. Cette technologie, faisant appel à des méthodes différentes, indépendantes et complémentaires au séquençage peut apporter une aide précieuse à l'assemblage des séquences. Par ailleurs, les cartes optiques ayant un pouvoir discriminant élevé, et permettant d'accéder au positionnement et à l'ordre des fragments de restriction dans le génome peuvent se révéler fort utiles lors de l'étude de clones épidémiques. Enfin, la possibilité offerte par les cartes optiques d'effectuer des comparaisons à grande échelle de l'architecture des génomes, en fait un outil particulièrement intéressant pour la mise en évidence de nouveaux réarrangements génétiques.

Remerciements

Nous tenons à remercier le directeur de recherche Benoit Cournoyer (Responsable de l'équipe Bactéries pathogènes opportunistes et environnement, Laboratoire d'écologie microbienne, Université Lyon 1, UMR5557, Université de Lyon, CNRS, VetAgroSup) et le professeur Éric Oswald (Laboratoire de Bactériologie-Hygiène, CHU Toulouse, Faculté de médecine Toulouse-Purpan) pour leur relecture attentive et leurs remarques constructives.

Références bibliographiques

- Aston C, Mishra B, Schwartz DC. 1999. Optical mapping and its potential for large-scale sequencing projects. *Trends Biotechnol* 17: 297-302.
- Barrangou R, Briczinski EP, Traeger LL, Loquasto JR, Richards M, Horvath P, Coute-Monvoisin A-C, Leyer G, Rendulic S, Steele JL, Broadbent JR, Oberg T, Dudley EG, Schuster S, Romero DA, Roberts RF 2009. Comparison of the Complete Genome Sequences of *Bifidobacterium animalis* subsp. *lactis* DSM 10140 and BI-04. *J. Bacteriol.* 191: 4144-4151.
- Cai W, Jing J, Irvin B, Ohler L, Rose E, Shizuya H, Kim UJ, Simon M, Anantharaman T, Mishra B, Schwartz DC. 1998. High-resolution restriction maps of bacterial artificial chromosomes constructed by optical mapping. *Proc Natl Acad Sci U S A* 95: 3390-3395.
- Chen Q, Savarino SJ, Venkatesan MM. 2006. Subtractive hybridization and optical mapping of the enterotoxigenic *Escherichia coli* H10407 chromosome: isolation of unique sequences and demonstration of significant similarity to the chromosome of *E. coli* K-12. *Microbiology* 152: 1041-1054.
- Foley SL, Simjee S, Meng J, White DG, McDermott PF, Zhao S. 2004. Evaluation of Molecular Typing Methods for *Escherichia coli* O157:H7 Isolates from Cattle, Food, and Humans. *Journal of Food Protection* 67: 651-657.
- Foley SL, White DG, McDermott PF, Walker RD, Rhodes B, Fedorka-Cray PJ, Simjee S, Zhao S. 2006. Comparison of Subtyping Methods for Differentiating *Salmonella enterica* Serovar Typhimurium Isolates Obtained from Food Animal Sources. *J Clin Microbiol*: 3569-3577.
- Fujisawa T, Narikawa R, Okamoto S, Ehira S, Yoshimura H, Suzuki I, Masuda T, Mochimaru M, Takaichi S, Awai K, Sekine M, Horikawa H, Yashiro I, Omata S, Takarada H, Katano Y, Kosugi H, Tanikawa S, Ohmori K, Sato N, Ikeuchi M, Fujita N, Ohmori M. 2010. Genomic Structure of an Economically Important Cyanobacterium, *Arthrospira (Spirulina) platensis* NIES-39. *DNA Research* 17: 85-103.
- Giongo A, Tyler HL, Zipperer UN, Triplett EW. 2010. Two genome sequences of the same bacterial strain, *Gluconacetobacter diazotrophicus* PAI 5, suggest a new standard in genome sequence submission. *Standards in Genomic Sciences*; Vol 2, No 3 (2010): Standards in Genomic Sciences.
- Hayashi T, Makino K, Ohnishi M, Kurokawa K, Ishii K, Yokoyama K, Han C-G, Ohtsubo E, Nakayama K, Murata T, Tanaka M, Tobe T, Iida T, Takami H, Honda T, Sasakawa C, Ogasawara N, Yasunaga T, Kuhara S, Shiba T, Hattori M, Shinagawa H. 2001. Complete Genome Sequence of Enterohemorrhagic *Escherichia coli* O157:H7 and Genomic Comparison with a Laboratory Strain K-12. *DNA Research* 8: 11-22.
- Jing J, Lai Z, Aston C, Lin J, Carucci DJ, Gardner MJ, Mishra B, Anantharaman TS, Tettelin H, Cummings LM, Hoffman SL, Venter JC, Schwartz DC. 1999. Optical mapping of *Plasmodium falciparum* chromosome 2. *Genome Res* 9: 175-181.
- Johnson PDR, Ballard SA, Grabsch EA, Stinear TP, Seemann T, Young HL, Grayson ML, Howden BP. 2010. A Sustained Hospital Outbreak of Vancomycin-Resistant *Enterococcus faecium* Bacteremia due to Emergence of vanB *E. faecium* Sequence Type 203. *Journal of Infectious Diseases* 202: 1278-1286.
- Karama M, Gyles CL. 2010. Methods for Genotyping Verotoxin-Producing *Escherichia coli*. *Zoonoses and Public Health* 57: 447-462.
- Kotewicz ML, Jackson SA, LeClerc JE, Cebula TA. 2007. Optical maps distinguish individual strains of *Escherichia coli* O157:H7. *Microbiology* 153: 1720-1733.
- Kotewicz ML, Mammel MK, LeClerc JE, Cebula TA. 2008. Optical mapping and 454 sequencing of *Escherichia coli* O157:H7 isolates linked to the US 2006 spinach-associated outbreak. *Microbiology* 154: 3518-3528.
- Lai Z, Jing J, Aston C, Clarke V, Apodaca J, Dimalanta ET, Carucci DJ, Gardner MJ, Mishra B, Anantharaman TS, Paxia S, Hoffman SL, Craig Venter J, Huff EJ, Schwartz DC. 1999. A shotgun optical map of the entire *Plasmodium falciparum* genome. *Nat Genet* 23: 309 - 313.



Méthodes

Latreille P, Norton S, Goldman B, Henkhaus J, Miller N, Barbazuk B, Bode H, Darby C, Du Z, Forst S, Gaudriault S, Goodner B, Goodrich-Blair H, Slater S. 2007. Optical mapping as a routine tool for bacterial genome sequence finishing. *BMC Genomics* 8: 321.

Lim A, Dimalanta ET, Potamouisis KD, Yen G, Apodaca J, Tao C, Lin J, Qi R, Skiadas J, Ramanathan A, Perna NT, Plunkett G, Burland V, Mau B, Hackett J, Blattner FR, Anantharaman TS, Mishra B, Schwartz DC. 2001. Shotgun optical maps of the whole *Escherichia coli* O157:H7 genome. *Genome Res* 11: 1584-1593.

Machida M, Asai K, Sano M, Tanaka T, Kumagai T, Terai G, Kusumoto K-I, Arima T, Akita O, Kashiwagi Y, Abe K, Gomi K, Horiuchi H, Kitamoto K, Kobayashi T, Takeuchi M, Denning DW, Galagan JE, Nierman WC, Yu J, Archer DB, Bennett JW, Bhatnagar D, Cleveland TE, Fedorova ND, Gotoh O, Horikawa H, Hosoyama A, Ichinomiya M, Igarashi R, Iwashita K, Juvvadi PR, Kato M, Kato Y, Kin T, Kokubun A, Maeda H, Maeyama N, Maruyama J, Nagasaki H, Nakajima T, Oda K, Okada K, Paulsen I, Sakamoto K, Sawano T, Takahashi M, Takase K, Terabayashi Y, Wortman JR, Yamada O, Yamagata Y, Anazawa H, Hata Y, Koide Y, Komori T, Koyama Y, Minetoki T, Suharnan S, Tanaka A, Isono K, Kuhara S, Ogasawara N, Kikuchi H. 2005. Genome sequencing and analysis of *Aspergillus oryzae*. *Nature* 438: 1157-1161.

Musmann M, Hu FZ, Richter M, de Beer D, Preisler A, Jørgensen BB, Huntemann M, Glöckner FO, Amann R, Koopman WJH, Lasken RS, Janto B, Hogg J, Stoodley P, Boissy R, Ehrlich GD. 2007. Insights into the Genome of Large Sulfur Bacteria Revealed by Analysis of Single Filaments. *PLoS Biol* 5: e230.

Neely RK, Deen J, Hofkens J. 2010. Optical mapping of DNA: Single-molecule-based methods for mapping genomes. *Biopolymers*: n/a-n/a

Neto M, Thevenot D, Loukiadis E. En préparation. O157:H7 and O26:H11 STEC from humans and animal origins show genomic diversity but share a common profile of virulence associated mobile element.

Nierman WC, Pain A, Anderson MJ, Wortman JR, Kim HS, Arroyo J, Berriman M, Abe K, Archer DB, Bermejo C, Bennett J, Bowyer P, Chen D, Collins M, Coulsen R, Davies R, Dyer PS, Farman M, Fedorova N, Fedorova N, Feldblyum TV, Fischer R, Fosker N, Fraser A, Garcia JL, Garcia MJ, Goble A, Goldman GH, Gomi K, Griffith-Jones S, Gwilliam R, Haas B, Haas H, Harris D, Horiuchi H, Huang J, Humphray S, Jimenez J, Keller N, Khouri H, Kitamoto K, Kobayashi T, Konzack S, Kulkarni R, Kumagai T, Lafton A, Latge J-P, Li W, Lord A, Lu C, Majoros WH, May GS, Miller BL, Mohamoud Y, Molina M, Monod M, Mouyna I, Mulligan S, Murphy L, O'Neil S, Paulsen I, Penalva MA, Perteau M, Price C, Pritchard BL, Quail MA, Rabbinowitsch E, Rawlins N, Rajandream M-A, Reichard U, Renauld H, Robson GD, de Cordoba SR, Rodriguez-Pena JM, Ronning CM, Rutter S, Salzberg SL, Sanchez M, Sanchez-Ferrero JC, Saunders D, Seeger K, Squares R, Squares S, Takeuchi M, Tekaia F, Turner G, de Aldana CRV, Weidman J, White O, Woodward J, Yu J-H, Fraser C, Galagan JE, Asai K, Machida M, Hall N, Barrell B, Denning DW. 2005. Genomic sequence of the pathogenic and allergenic filamentous fungus *Aspergillus fumigatus*. *Nature* 438: 1151-1156.

Northey G, Gal M, Rahmati A, Brazier JS. 2005. Subtyping of *Clostridium difficile* PCR ribotype 001 by REP-PCR and PFGE. *J Med Microbiol* 54: 543-547.

Ogura Y, Ooka T, Iguchi A, Toh H, Asadulghani M, Oshima K, Kodama T, Abe H, Nakayama K, Kurokawa K, Tobe T, Hattori M, Hayashi T. 2009. Comparative genomics reveal the mechanism of the parallel evolution of O157 and non-O157 enterohemorrhagic *Escherichia coli*. *Proceedings of the National Academy of Sciences* 106: 17939-17944.

Perna NT, Plunkett G, Burland V, Mau B, Glasner JD, Rose DJ, Mayhew GF, Evans PS, Gregor J, Kirkpatrick HA, Posfai G, Hackett J, Klink S, Boutin A, Shao Y, Miller L, Grotbeck EJ, Davis NW, Lim A, Dimalanta ET, Potamouisis KD, Apodaca J, Anantharaman TS, Lin J, Yen G, Schwartz DC, Welch RA, Blattner FR. 2001. Genome sequence of enterohaemorrhagic *Escherichia coli* O157:H7. *Nature* 409: 529-533.

Saunders, Wu, Abuoun, Pan, Anjum, Woodward. 2010. Optical genetic mapping defines regions of chromosomal variation in serovars of *S. enterica* subsp. *enterica* of concern for human and animal health. *Epidemiology and Infection*.

Schwan WR, Briska A, Stahl B, Wagner TK, Zentz E, Henkhaus J, Lovrich SD, Agger WA, Callister, SM, DuChateau B, Dykes CW. 2010. Use of optical mapping to sort uropathogenic *Escherichia coli* strains into distinct subgroups. *Microbiology* 156: 2124-2135.

Shukla SK, Kislow J, Briska A, Henkhaus J, Dykes C. 2009. Optical mapping reveals a large genetic inversion between two methicillin-resistant *Staphylococcus aureus* strains. *J Bacteriol* 191: 5717-5723.

van het Hoog M, Rast T, Martchenko M, Grindle S, Dignard D, Hogues H, Cuomo C, Berriman M, Scherer S, Magee BB, Whiteway M, Chibana H, Nantel A, Magee PT. 2007. Assembly of the *Candida albicans* genome into sixteen supercontigs aligned on the eight chromosomes. *Genome Biology* 8: R52.

Zhou S, Deng W, Anantharaman TS, Lim A, Dimalanta ET, Wang J, Wu T, Chunhong T, Creighton R, Kile A, Kvikstad E, Bechner M, Yen G, Garic-Stankovic A, Severin J, Forrest D, Runnheim R, Churas C, Lamers C, Perna NT, Burland V, Blattner FR, Mishra B, Schwartz DC. 2002. A whole-genome shotgun optical map of *Yersinia pestis* strain KIM. *Appl Environ Microbiol* 68: 6321-6331.

Zhou S, Kile A, Bechner M, Place M, Kvikstad E, Deng W, Wei J, Severin J, Runnheim R, Churas C, Forrest D, Dimalanta ET, Lamers C, Burland V, Blattner FR, Schwartz DC. 2004a. Single-molecule approach to bacterial genomic comparisons via optical mapping. *J Bacteriol* 186: 7773-7782.

Zhou S, Kile A, Kvikstad E, Bechner M, Severin J, Forrest D, Runnheim R, Churas C, Anantharaman, TS, Myler P, Vogt C, Ivens A, Stuart K, Schwartz DC. 2004b. Shotgun optical mapping of the entire *Leishmania* major Friedlin genome. *Mol Biochem Parasitol* 138: 97-106.