EuroReference
journal of Reference

anses
French agency for food, environmental and occupational health & safety

Journal No. 5
Summer 2011

Methods

Optical maps: methodology and applications in microbiology

Neto M. (1), Skorski G. (2), Thevenot D. (1), Loukiadis E. (1)
(1) VetAgroSup Campus vétérinaire de Lyon, Unité CALITYSS - Equipe EMSA-Lyon, LMAP - LNR STEC. 1 avenue Bourgelat. 69280 Marcy l'Etoile, France.
(2) PHYLOGENE, 62 RN 113. 30620 Bernis, France.

Optical maps provide a graphical representation of the location of restriction sites in the whole genome of organisms under study. They can be obtained semi-automatically using a method developed by the company OpGen. Optical maps offer many advantages when undertaking and completing the sequencing of both prokaryotic and eukaryotic genomes. In epidemiology, optical mapping is one of the few tools available for studying the whole genome. In addition, clones of interest can be characterised in great detail, and epidemic clones can be differentiated from other clonal lines. Finally, in comparative genomics studies, optical maps can be used to quickly and easily compare the genomes of several bacteria and identify gene insertions/deletions or even as yet unidentified rearrangements.

Introduction

Optical maps provide a graphical representation of the location of restriction sites in the whole genome of organisms under study. After an overview on how optical maps are constructed, a "mini-review" of the literature on applications of these maps, e.g. in genome sequencing, epidemiology and comparative genomics, will be presented. Each of these applications will be explained and illustrated through examples found in the literature. It should be noted that the use of optical maps for studying non-culturable microorganisms or identifying organisms in complex samples (clinical samples, etc.) has also been reported (Mussmann *et al.* 2007; Skorski, personal communication).

Materials and methods

Methodology

The general methodology of optical maps is based on mapping the location of restriction sites in the whole genome of the organism studied (Figure 1-1).
OpGen is the only company to have developed a semi-automated method which performs all the steps required to obtain an optical map from chromosomal DNA (Figure 1-2).
Chromosomal DNA is extracted from the organisms studied using a 'soft' method (for example using magnetic beads) to obtain large fragments of DNA (Figure 1-1A). These strands are then immobilised on a special microscope slide (Figure 1-1B). After verifying its integrity, its concentration (10^5 bacterial genomes on average are needed to construct a map) and the absence of DNA strands <150 kb, the immobilised DNA is then digested with a restriction enzyme adapted to the organism and purpose of the study (for the same organism several enzymes can be used independently) (Figure 1-1C). The restriction enzyme used is selected using software (such as Enzyme Chooser by OpGen) in order to obtain an average fragment size of between 6 and 12 kb and no restriction fragments larger than 80 kb. The restriction enzyme conventionally used for *Escherichia coli* is *Nco*1. After labelling with a fluorochrome (Figure 1-1D), the DNA fragments are visualised under a fluorescence microscope coupled with analytical software. The size of the fragments is measured (by comparison with a size standard) and the order of fragments is taken into account (Figure 1-1D).
Fragments smaller than 600 bp are not detected. The individual restriction maps are then assembled by matching fragment profiles to construct the optical map of the whole genome of the bacteria under study (Figure 1-1E). The resulting restriction map of the whole genome resembles a bar code (Figure 1-1F), with each bar corresponding to a restriction site. The gap between each bar corresponds to the number of base pairs found between the two restriction sites. The maps are then analysed by computer-assisted interpretation software such as MapSolver™.

Limits

One current limitation of this technique is its availability, as access to this semi-automated technology is only possible *via* the company OpGen.
Moreover, the cost of producing a map is about €380 (price reported by OpGen), the automatic system and integrated software cost around €260K and maintenance costs are approximately €21K per year.
The optical mapping technique is unable to analyse DNA strands smaller than about 150 kb (such as certain plasmids), as these strands do not contain enough restriction sites to perform a reliable alignment.
The effectiveness and limitations of optical maps, particularly with regard to detection of small fragments, have been tested by comparing restriction maps obtained *in silico* from sequenced strains (hereinafter called *in silico* maps) with actual optical maps of these strains (including *E. coli* O157:H7 (Kotewicz *et al.,* 2007) and *Yersinia pestis* (Zhou *et al.,* 2002)).
For example, comparison of optical maps and *in silico* maps of two sequenced strains of *E. coli* O157:H7 (EDL933 and Sakai) showed that out of the 639 *in silico* fragments obtained for the Sakai strain, only 96 were not detected by the optical map and all were smaller than 2 kb (the majority of these - 60/96 - were also less than 600 bp because these fragments are too small to be retained on the surface of the slide). These undetected fragments therefore only accounted for 20,309 bp out of a total of 5.5 Mb (approximately 0.37%) (Kotewicz *et al.,* 2007). These studies also showed that the 12 differences detected in the EDL933 strain between the *in silico* map and the optical map were due to polymorphisms at the *Bam*HI restriction sites or absence of cleavage at six positions. These

EuroReference
journal of Reference
anses
French agency for food, environmental and occupational health & safety

Journal No. 5
Summer 2011

Methods

differences could thus be attributed either to the technological limitations of the maps (the presence of small undetected fragments) or to polymorphisms in the genome of the studied strain. Nevertheless, the authors conclude that the optical map methodology is an accurate and useful tool for differentiating and studying highly clonal strains.

Support for the sequencing of whole genomes

Sequence assembly accounts for a large proportion of the cost and time involved in sequencing. Optical maps are increasingly used to shorten timeframes and increase the reliability of sequence assembly for whole genomes of both prokaryotic and eukaryotic organisms. Table 1 lists examples from the literature.

Constructing optical maps of a microorganism being sequenced serves i) to determine the size of the genome; ii) as the basis for sequence assembly; iii) to confirm, by comparing the optical map constructed on the DNA of the organism and the map produced *in silico*, the correct alignment and/or to highlight any gaps in the final sequence (Figure 2 illustrates schematically the support provided by optical maps in the resolution and verification of assemblies from the sequencing of a whole genome (modified from Aston *et al.,* (1999) and Neely *et al.* (2010)). The *in silico* map of the contiguous sequence obtained from sequencing is compared to the optical map of the bacterial strain studied. This comparison confirms the assembly and direction of contiguous sequences, highlights any gaps

Figure 1. Method developed for microorganisms by OpGen for constructing an optical map (with permission of OpGen)

1. Method:
A- DNA is extracted from bacterial cells using a magnetic bead method in order to obtain large fragments of DNA. B- DNA is attached to a specific support. C- DNA is digested with an appropriate restriction enzyme (red arrows indicate restriction sites) and labelled with a fluorochrome. D- The fluorescence intensity is measured to determine the size of the fragments and their order is recorded. E- Overlapping DNA strands are assembled to produce an optical map of the complete genome. F- The optical map is shown here, each vertical line represents the location of a restriction site and the space between the bars represents the size in kilobases of the restriction fragments obtained;

2. Automated method developed for microorganisms by OpGen:
A- The DNA is deposited on the Argus MapCard™ cartridge, quantified and digested in the Argus™ Optical Mapping System which creates a map. B- The generated map is then read and analysed automatically by the Argus™ Optical Mapper.

Methods

and helps in their resolution); iv) to verify the sequencing of genomes with many prophage insertions or regions featuring repeated sequences that are often sources of mismatches with traditional tools (Jing *et al.* 1999; Lim *et al.* 2001).
Producing two optical maps with two restriction enzymes, generating results from two screens and highlighting the location and distribution of two types of restriction sites, may also provide additional information and assistance with sequence assembly (this approach was used for sequencing Yersinia pestis by Zhou *et al.* (2002)).
In addition, as optical maps only require a small amount of DNA (105 bacterial genomes on average), they can be used for organisms which are difficult to culture and clone such as *Plasmodium falciparum* (Jing *et al.* 1999).
Furthermore, compared to other sequence assembly support tools such as restriction maps and hybridisation, optical maps have the advantage of not requiring steps that can lead to artefacts (such as the bias associated with cloning which is sometimes needed before sequencing) and can be performed in parallel and independently of sequencing.
Work performed by Giongo *et al.* (2010) on *Gluconacetobacter diazotrophicus* also particularly illustrates this point: two independent sequencing runs of the same PAI5 strain revealed differences and the optical maps confirmed one of the PAI5 strain sequences.
Thus, optical maps seem to be a reliable and extremely useful tool for assisting with or verifying assembly of a sequenced genome (Giongo *et al.*, 2010; Zhou *et al.*, 2004b) particularly as they use a principle and methods that are different, independent and complementary to sequencing.

A tool for typing and comparing microbial strains that can be used in epidemiology

Many tools are now available to study the clonal relationships among strains isolated during epidemics or poly-infections: pulsed-field gel electrophoresis (PFGE); multilocus sequence typing (MLST); multiple-locus variable-number tandem-repeats analysis (MLVA); repetitive-sequence-based PCR (rep-PCR), etc. Among them, only three can be used to study the whole genome: PFGE, rep-PCR and optical mapping.
rep-PCR appears to have the same discriminatory power as PFGE for some organisms (such as *Salmonella* enterica (Foley *et al.* 2006) and *Clostridium difficile* (Northey *et al.* 2005) but to be less discriminating for other organisms such as *E. coli* O157:H7 (Foley *et al.* 2004). The tools available and applicable to Shiga toxin-producing *E. coli* (STEC) were compared in a recent review (Karama and Gyles, 2010). The authors also demonstrate that optical maps and PFGE have similar performance in terms of discriminatory power, ease of implementation and standardisation, but both involve difficulties in interpretation and laborious analysis of results. However, with regard to the time needed and complexity of analysis of data obtained by optical maps (compared to PFGE), the aims should be taken into account. Accordingly, if the aim is to study the clonal distances between strains, analysis of data obtained by optical maps is no more time-consuming nor more complex than that obtained by PFGE. Moreover, if the aim is to study the characteristics of epidemic clones or groups of strains (which is impractical with PFGE), optical maps seem to be more discriminating, due to the greater number of fragments obtained and analysed by optical mapping compared to PFGE. Kotewicz

Table 1. Examples of the use of optical maps as an aid to the sequencing of complete genomes.

Organisms	Uses	References
Large genomes and eukaryotes		
Aspergillus oryzae	Verification of the sequence	(Machida *et al.*, 2005)
Aspergillus fumigatus	Assistance to assembly	(Nierman *et al.*, 2005)
Candida albicans	Verification of the sequence	(van het Hoog *et al.*, 2007)
Human BAC library	Assistance to assembly	(Cai *et al.*, 1998)
Leishmania major Friedlin	Assistance with resolving contig gaps and errors	(Zhou *et al.*, 2004b)
Plasmodium falciparum (parasite)	Organism difficult to culture, little DNA required to generate the optical map; assistance with assembly	(Jing *et al.*, 1999)
	Assistance to assembly	(Lai *et al.*, 1999)
Prokaryotes		
Arthrospira platensis	Assistance to assembly	(Fujisawa *et al.*, 2010)
Large Sulfur Bacteria (*Beggiatoa*)	Assistance to analysis of non-culturable bacteria	(Mussmann *et al.*, 2007)
Escherichia coli O157:H7	Assistance to analysis of contigs *via* information on rearrangements	(Lim *et al.*, 2001)
Gluconacetobacter diazotrophicus PAI 5	Comparison of sequencing of the same strain by 2 teams. Shows the errors that may arise with rapid assemblies and/or rearrangements and the benefit of optical mapping in proposing new sequences.	(Giongo *et al.*, 2010)
Xenorhabdus nematophila and *Xenorhabdus bovienii*	Assistance to assembly	(Latreille *et al.*, 2007)
Yersinia pestis	Use of 2 different enzymes: one to assist with assembly and the other for verification of the sequence. Highlighting of assembly errors and help with correct assembly *via* the optical map.	(Zhou *et al.*, 2002)

Methods

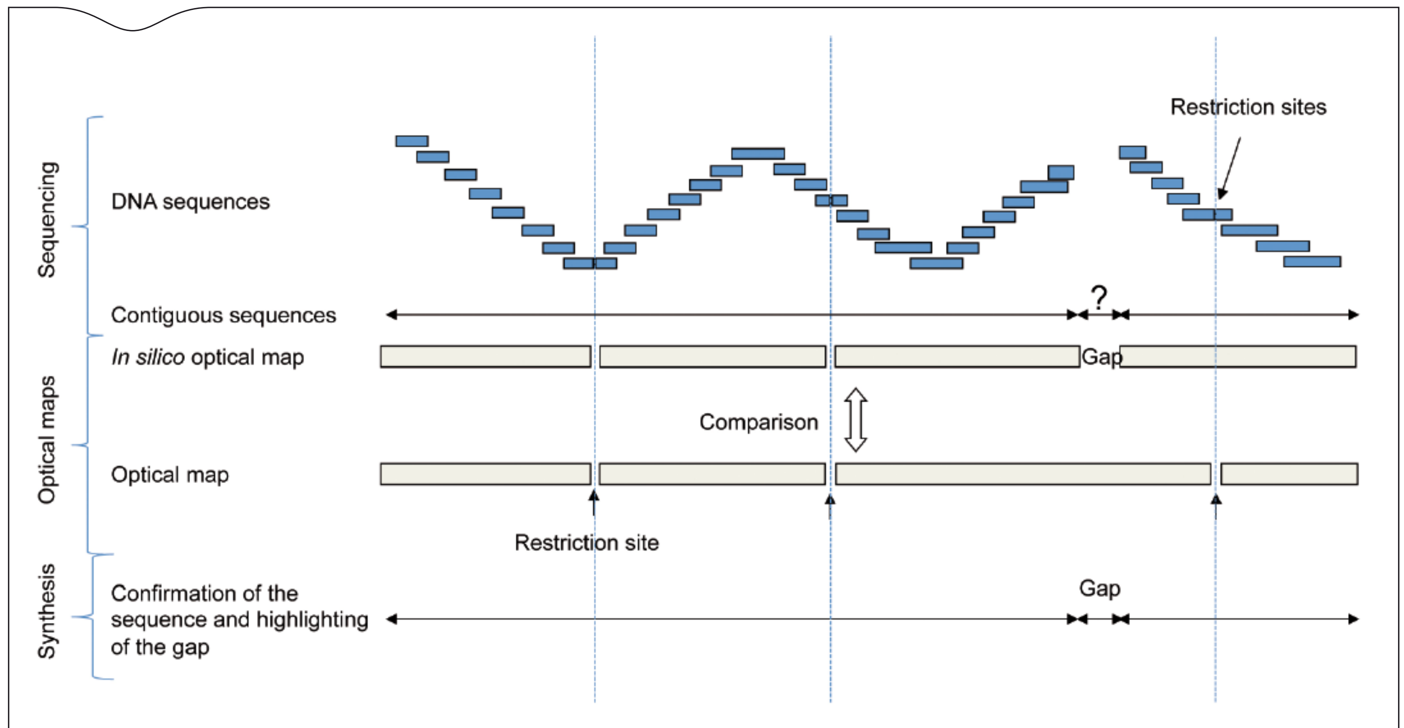

Figure 2. Schematic representation of the assistance provided by optical maps in the resolution and verification of assemblies of complete genome sequencing runs (modified from Aston *et al.* (1999) and Neely *et al.* (2010)).
The in silico *map of the contiguous sequence obtained from sequencing is compared to the optical map of the bacterial strain under study. This comparison makes it possible to i) confirm the assembly and direction of contiguous sequences, ii) highlight gaps and help resolve them. The optical maps can also serve as a scaffold in the assembly of sequences in order to obtain the contiguous sequence.*

et al. (2007) show that optical maps have sufficient resolution and detail to differentiate very similar strains of *E. coli* O157:H7. Comparing the number of fragments obtained by PFGE with the enzyme *Xba*I conventionally used for *E. coli* O157:H7 (43 fragments) to that obtained by optical mapping with the enzyme *Bam*HI (645 fragments) suggests that optical maps are more informative.

In the case of the sequenced 11368 strain of *E. coli* O26:H11 (Figure 3), the number of restriction sites obtained with the enzyme *Nco*I used for constructing optical maps (725 fragments, Figure 3A) was higher than that obtained with the enzyme *Xba*I used conventionally in PFGE (36 fragments, Figure 3B), as the choice of restriction enzyme for the latter is in fact limited by the visualisation of distinct fragments. Considering the number of fragments obtained in this particular example, the results from optical mapping provided a level of detail 20 times greater than that obtained by PFGE.

In the laboratory, eight strains of foodborne *E. coli* O26:H11 and the sequenced O26 strain 11368 of human origin have been typed and compared by PFGE (Figure 3D) and by optical mapping (Figure 3E). The two main clusters obtained by these two techniques were similar and supported the conclusion that strains 108.1.1, 60.1 and 64.1 belong to the same clone and form a separate cluster from the other strains of *E. coli* O26:H11 studied. However, the clonal relationships between the strains in cluster 1 (Figure 3) differed according to the two techniques, which may illustrate the greater discriminatory power of optical mapping compared to PFGE.

Finally, compared with results obtained by PFGE, optical maps have the advantage of ordering the fragments according to genome and not according to fragment size (Figure 3B and 3C) and indicating the location of restriction fragments on the genome, which may be particularly useful for the characterisation of epidemic clones (Kotewicz *et al.* 2008) or groups of strains (Schwan *et al.* 2010).

Accordingly, by constructing optical maps of isolates of *E. coli* O157:H7 from contaminated spinach responsible for a foodborne disease outbreak in the USA, Kotewicz *et al.* (2008) identified 14 chromosome markers associated with an epidemic clone mainly by comparing *in silico* maps of sequenced EDL933 and Sakai genomes. By combining sequencing analyses and optical maps, the authors show that the epidemic clone implicated lacked the gene *stx*1 although it had a similar prophage to the prophage V of EDL933 that classically carries this gene. In addition, analysis of the optical maps, coupled with sequencing of potentially interesting chromosome markers, helped locate the *stx*2 and *stx*2c genes in the genome. These results suggest that the particular virulence of some strains of *E. coli* O157:H7 may well depend on the location or rearrangements of specific phage sequences.

Another study on uropathogenic *E. coli* (UPEC) (Schwan *et al.* 2010) has shown that, through the use of optical maps, two groups of UPECs could be differentiated. These two groups corresponded to profiles of presence/absence of virulence factors and antibiotic resistance observed among strains. The optical maps were presented by the authors as a possible tool

Methods

for the rapid detection of antibiotic-resistant strains during outbreaks.
Other examples of epidemiological studies using optical maps (recorded between 2006 and 2010) are listed in Table 2 and all emphasise the value of using them for the precise study of groups of strains and epidemic clones.
Moreover, as optical maps can also be used to reconstruct genomes by aligning strands of DNA (with a limit of detection of about 105 bacteria), they can be used for the study of non-culturable organisms (Mussmann *et al.* 2007) as well as for the study of complex polymicrobial samples (clinical samples, etc.) in order to identify new strains or even to study cooperation between strains (Skorski, personal communication).

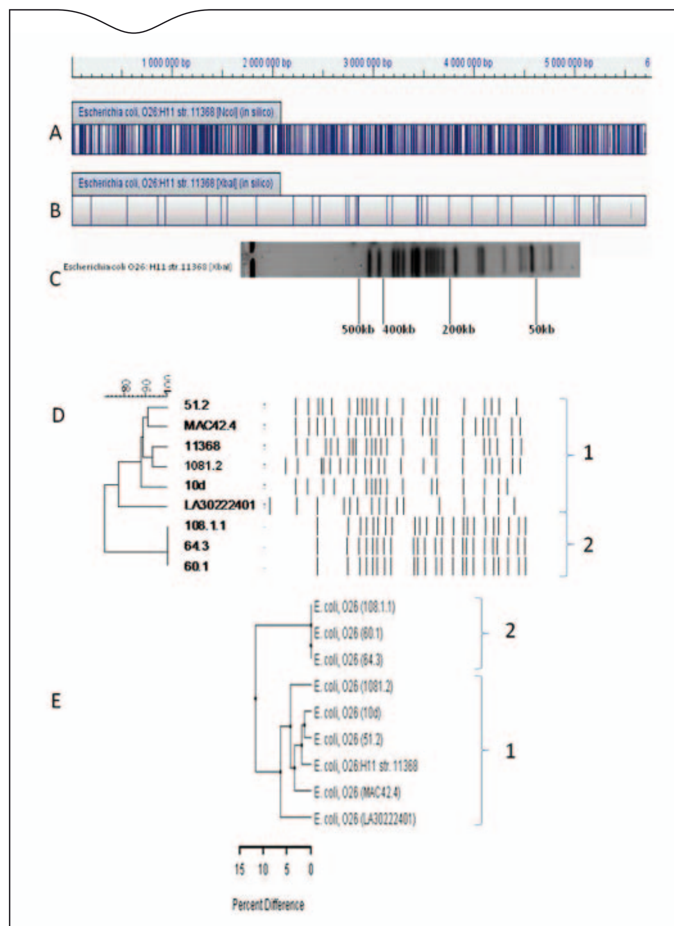

Figure 3. Optical maps, PFGE and similarity clusters for strains of *E. coli* O26:H11.
A- In silico optical map of the sequenced E. coli *O26:H11 strain 11368 isolated from a patient with diarrhoea (Ogura* et al.*, 2009) using the enzyme* NcoI *(725 fragments); B- In silico optical map of the* E. coli *O26:H11 strain 11368 using the enzyme* XbaI *(36 fragments); C- PFGE obtained using the enzyme* XbaI *from the* E. coli *O26:H11 strain 11368. D- Similarity clusters obtained by PFGE from nine strains of* E. coli *O26:H11 of human or food origin. Restriction using the enzyme* XbaI*. Dendrogram generated with GelComparII (Applied-Maths) using the Dice coefficient with a tolerance of 1.5% on the position of the bands and UPGMA (unweighted pair group method with arithmetic mean), the scale represents the percentage of similarity. E- Similarity clusters obtained from the analysis of optical maps of nine strains of* E. coli *O26:H11 of human or food origin. Dendrogram generated with optical maps (with the enzyme* NcoI*) using UPGMA, the scale represents the percentage of difference.*

There are a limited number of typing tools available for the study of whole genomes (PFGE, rep-PCR, optical maps) and each has its advantages and disadvantages in terms of feasibility, availability, discriminatory power and cost. It is therefore necessary to adapt the tool used to the objective of the study, particularly with regard to the accuracy of the desired results, as well as the timeframe for obtaining results and the budget available. Although optical mapping is less available (technology only available through OpGen) and more expensive than tools such as PFGE and rep-PCR, it is more informative since it is more discriminating than PFGE (20 times more fragments obtained, for example, for the 11368 strain of *E. coli* O26:H11) and above all it provides information on the positioning and order of restriction fragments in the genome. These characteristics make it a very accurate and useful tool, especially for the detailed analysis of epidemic clones.

A tool for comparative genomics of the whole genome: detection of insertions, deletions, inversions and translocations

Tools for typing, PCR and sequencing of complete bacterial genomes have helped highlight the great genomic plasticity (rearrangements, deletions, insertions) of microorganisms, which explains their diversity and evolution. However, optical maps may have advantages over these conventional comparative genomics tools:

- DNA microarrays and hybridisation can screen a large number of genes, quickly and on a large number of strains. However, unlike optical mapping, they can only detect the presence or absence of genes and cannot study unknown genes or potential gene insertions and rearrangements;
- typing by macro-restriction (such as PFGE) highlights differences between genomes, but unlike optical mapping, cannot directly link these differences to genes or to a position in the genome;
- "Shotgun" or high-throughput sequencing has the advantage of being more comprehensive because it provides more

Table 2. Examples of the use of optical maps in epidemiological studies

Organisms	Uses	References
Enterotoxigenic *E. coli* (ETEC)	Comparison of the *E. coli* K12 strain with the ETEC strain H10407. Similar strains.	(Chen *et al.*, 2006)
Uropathogenic *E. coli* (UPEC)	Comparison between UPEC, *Shigella*, EHEC and the K12 strain. UPEC form a separate group. Among the UPEC, 2 groups correlated with the presence/absence of virulence and antimicrobial resistance factors.	(Schwan *et al.*, 2010)
E. coli O157:H7	Inter-comparison of strains and comparative genomics study (see part 4 below).	(Kotewicz *et al.*, 2007)
E. coli O157:H7	Study of an epidemic clone. Highlighting of 14 specific chromosome markers (see below).	(Kotewicz *et al.*, 2008)

Methods

information (the complete genome sequence can be obtained) but, to date, optical mapping remains the most suitable tool for studying species for which few sequences are available.
Kotewicz *et al.* (2007) and Zhou *et al.* (2004a) have shown the speed and ease of analysis of optical maps, mainly due to their visual and graphical aspect. Moreover, the correspondence between areas of interest (identified during the analysis of maps) and gene annotation can be demonstrated through comparison with an *in silico* map of a sequenced strain (Figure 4).
Nevertheless, while optical maps can highlight new rearrangements and new genetic loci, the sequencing of the region exhibiting these new rearrangements or loci is still necessary to confirm the observations made with optical mapping (Shukla *et al.,* 2009).
Furthermore, it should be noted that the level of resolution of optical maps is insufficient for highlighting point or small mutations.
Several studies have shown how optical maps can contribute to comparative genomics.
The comparison of optical maps of the sequenced EDL933 and Sakai strains of *E. coli* O157:H7 only showed two significant differences relating to the size of some fragments and the existence of an inversion of 430kb (or 7.8% of the genome) in a region flanked by prophages O and P in EDL933 (Kotewicz *et al.,* 2007).
Optical maps of the EDL933 and Sakai strains have also been compared to nine other strains of *E. coli* O157:H7 (Kotewicz *et al.,* 2007). In total, 91 differences were identified; all clustered in 28 loci of which the majority corresponded to prophages, cryptic prophages or possibly new phage insertion sites. This study also shows that the two most variable sites were those occupied by prophages carrying the *stx*1 and *stx*2 genes.
The comparison of optical maps of four serovars of *Salmonella* enterica subsp enterica (Saunders *et al.* 2010) highlighted variations in the genome of S. Typhimurium, associated with the presence of prophages and correlated with observed antibiotic resistance profiles, opening new topics of research.
A study underway in our laboratory (Neto *et al.* in preparation) aims to compare STEC O157:H7 strains of human or food origin using optical maps, in a collaboration with OpGen. Comparing optical maps of the human LS32 strain (LMAP-NRL STEC collection) and two sequenced human strains (EDL933 and Sakai) (Figure 5A) showed that these strains were very closely related in terms of genomic structure and that differences were mainly in an area rich in 'O-islands' and prophages (Figure 5B). In the case of a non-sequenced strain, optical maps help to quickly guide the search by PCR for the position of prophages in the genome. These results illustrate the large extent to which prophages contribute to the genome plasticity of *E. coli* O157:H7. Increasing the number and diversity of microorganisms studied by optical mapping (especially on a wider range of Enterobacteriaceae) could allow the detection of new areas of instability and mobile genetic elements involved in the emergence of *E. coli* pathogens.
Optical maps, due to their technical advantages (no prior information needed on the genome, access to information on the position and order of restriction fragments on the genome) can be used to study species for which few data are available, to highlight the presence of new genes and to locate new gene insertions and as yet unidentified rearrangements.

Conclusion

Optical maps have many applications, for which they have distinct advantages but also disadvantages relating to its very methodology, its availability (technology only available through OpGen) and its cost. These aspects are summarised for each application in Table 4.
As this technology uses methods that are different, independent and complementary to sequencing, it may provide valuable assistance in sequence assembly. Moreover, as optical maps have high discriminatory power, and provide access to the location and order of restriction fragments in the genome, they may be very useful in the study of epidemic clones. Lastly, the possibility offered by optical maps for conducting large-scale comparisons of genome architecture, makes it a particularly interesting tool for the identification of new genetic rearrangements.

Acknowledgements
We would like to thank Benoit Cournoyer, Senior Researcher

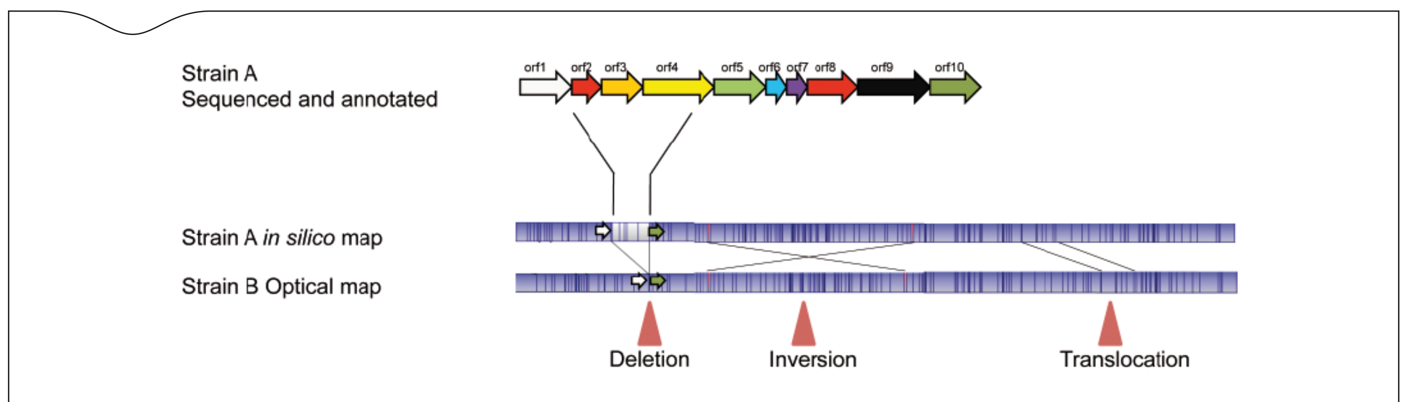

Figure 4. Schematic representation of the contribution of optical maps to comparative genomics (comparing a sequenced strain and a strain of interest).
The vertical bars in the optical maps represent the restriction sites, the arrows represent the ORFs. The optical map of strain A was produced in silico *from the nucleotide sequence available in databases. The combined use of optical maps and sequences can highlight differences between the genomes such as insertions, translocations, inversions and deletions (represented by triangles).*

Methods

Table 3. Examples of the use of optical maps in comparative genomics studies

Organisms	Uses	References
Bifidobacterium animalis	Comparison of strains. Few differences between strains.	(Barrangou *et al.*, 2009)
Enterococcus faecium	Comparison of strains and highlighting of differences between epidemic clones.	(Johnson *et al.*, 2010)
Enterotoxigenic *E. coli* (ETEC)	Comparison of strains. Demonstration of strong similarities between *E. coli* K12 and ETEC H10407	(Chen *et al.*, 2006)
E. coli O157:H7	Comparison of strains and highlighting of variable regions, insertions, deletions and DNA inversions between strains.	(Kotewicz *et al.*, 2007)
E. coli O157:H7	Highlighting of 14 specific chromosome markers for an epidemic clone.	(Kotewicz *et al.*, 2008)
Salmonella enterica subsp *enterica*	Comparison of strains. Demonstration that deletions of certain cassettes are responsible for various antibioresistance profiles. Highlighting of as yet uncharacterised genetic rearrangements.	(Saunders *et al.*, 2010)
Shigella flexneri	Comparison of strains. Highlighting of new sites bearing insertion sequences and phage-related gene insertions.	(Zhou *et al.*, 2004a)

Table 4. Advantages and disadvantages of the use of optical maps in the various fields described

Applications of optical maps	Advantages	Disadvantages
Assistance with and/or verification of sequencing of genomes	Independent of sequencing (method can be implemented in parallel) Does not require cloning or other techniques that can create artefacts Fast and reliable	Reduced availability* No information on plasmids
Epidemiology	Analysis based on the complete genome Strong discriminatory power (200 markers per Mb) Ease of implementation Ease of standardisation Information provided on the order of fragments in the genome and possible comparison with sequenced strains Information provided on the annotation of sequenced strains, and annotation of optical maps possible Includes cluster function	Interpretation not easy and analysis of results laborious Reduced availability* Cost** No information on plasmids
Comparative genomics	Speed and ease of analysis with the software Visual and graphical aspect Easy comparison of multiple strains Possibility of demonstrating new as yet uncharacterised rearrangements Information provided on the order of fragments in the genome and possible comparison with sequenced strains Possibility of adapting it to other techniques such as "DNA probing" (labelled probes) Information provided on the annotation of sequenced strains, and annotation of optical maps possible Includes cluster function	Reduced availability* Cost** No information on plasmids smaller than 150Kb

* The technology is only available from OpGen.
** The cost of each map is about €380, for the reagents. The cost of the automatic system and the software integrated in the machine amounts to approximately €260K with a €21K annual maintenance fee.

Methods

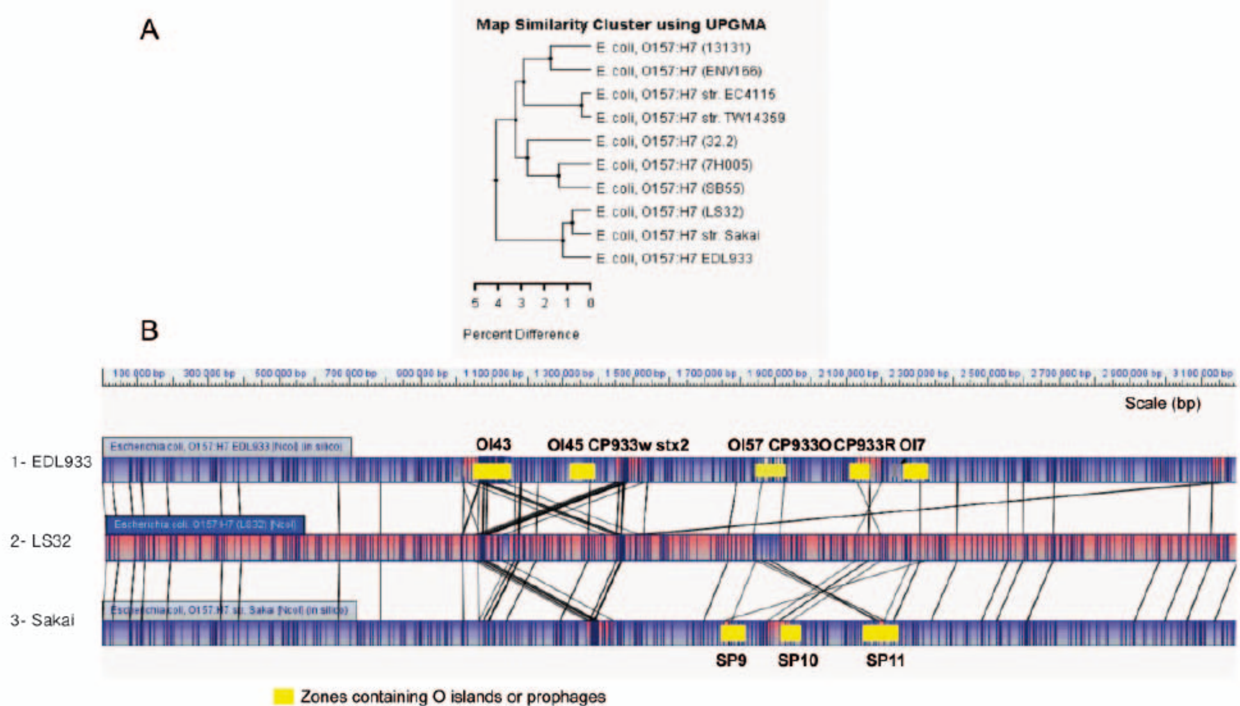

Figure 5. Comparison of strains of *E. coli* O157:H7.

A- Similarity clusters obtained from the analysis of optical maps of nine strains of E. coli *O157:H7 of human or food origin using the enzyme* NcoI*. Dendrogram generated with optical maps (with the enzyme* NcoI*) using UPGMA. The scale represents the percentage of difference. Optical maps for the Sakai, EDL933, TW14359 and EC4115 strains were produced in silico using Genbank sequences: NC_002695, NC_002655, NC_013008, NC_011353.*

B- Example of the most variable region identified by the comparison of three strains: Sakai, EDL933 and LS32. Optical maps were produced using the enzyme NcoI 1) in silico *map of the sequenced EDL933 strain, 2)* optical map of the LS32 strain, 3) *in silico map of the sequenced Sakai strain. The EDL933 and LS32 strains are divergent within the region corresponding to O Island 57 (CP933O) in the sequenced EDL933 strain. The region is different in the CP933R cryptic annotated prophage- region in EDL933 which has an inversion relative to the LS32 strain. There is an insertion in strain LS32 compared to the Sakai strain, this region is annotated in EDL933 as corresponding to O Island 43. The inversion shown by Kotewicz* et al. *(2007) between O islands 57 and 71, between the EDL933 and Sakai strains, is only found here in the Sakai strain. The LS32 strain has the same configuration as the EDL933 strain. This region has O Islands OI57 (CP933O), CP933R, OI68, OI69, OI70 and OI71 (Perna* et al.*, 2001). This region encodes prophages SP (Sakai prophage 9, SP10, SP11, SP12 (Hayashi* et al.*, 2001)).*

(Leader of the Opportunistic Bacterial Pathogens and Environment team, Laboratory of Microbial Ecology, University of Lyon 1, UMR5557, University of Lyon, CNRS, VetAgroSup) and Professor Eric Oswald (Laboratory of Bacteriology-Hygiene, University Hospital (CHU) of Toulouse, Faculty of Medicine Toulouse-Purpan) for their careful reading and constructive comments.

References

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 Bl-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.

Methods

Giongo A, Tyler HL, Zipperer UN, Triplett EW. 2010. Two genome sequences of the same bacterial strain, Gluconacetobacter diazotrophicus PAl 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.

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, Potamousis KD, Yen G, Apodoca 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.

Mussmann, 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. In preparation. 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, Pertea 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, Potamousis 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.