



Current strategies for mobilome research

Jørgensen, Tue Sparholt; Kiil, Anne S.; Hansen, Martin Asser; Sørensen, Søren Johannes; Hansen, Lars H.

Published in:
Frontiers in Microbiology

DOI:
[10.3389/fmicb.2014.00750](https://doi.org/10.3389/fmicb.2014.00750)

Publication date:
2015

Document version
Publisher's PDF, also known as Version of record

Citation for published version (APA):
Jørgensen, T. S., Kiil, A. S., Hansen, M. A., Sørensen, S. J., & Hansen, L. H. (2015). Current strategies for mobilome research. *Frontiers in Microbiology*, 5, [750]. <https://doi.org/10.3389/fmicb.2014.00750>



Current strategies for mobilome research

Tue S. Jørgensen^{1*}, Anne S. Kiil¹, Martin A. Hansen¹, Søren J. Sørensen^{1*} and Lars H. Hansen^{1,2*}

¹ Section of Microbiology, Department of Biology, University of Copenhagen, Copenhagen, Denmark

² Environmental Microbiology and Biotechnology, Department of Environmental Science, Aarhus University, Roskilde, Denmark

Edited by:

Holger Heuer, Julius Kühn-Institut – Federal Research Centre for Cultivated Plants, Germany

Reviewed by:

Jan Roelof Van Der Meer, University of Lausanne, Switzerland
Amparo Latorre, University of Valencia, Spain
Itzhak Mizrahi, Agricultural Research Organization – Volcani Research Center, Israel

*Correspondence:

Lars H. Hansen, Environmental Microbiology and Biotechnology, Department of Environmental Science, Aarhus University, Frederiksborgvej 399, 4000 Roskilde, Denmark
e-mail: lhha@envs.au.dk;
Søren J. Sørensen and Tue S. Jørgensen, Section for Microbiology, Department of Biology, University of Copenhagen, Universitetsparken 15, Building 1, 1st floor, Copenhagen 2100Kbh Ø, Denmark
e-mail: sjs@bio.ku.dk;
tuesparholt@gmail.com

Mobile genetic elements (MGEs) are pivotal for bacterial evolution and adaptation, allowing shuffling of genes even between distantly related bacterial species. The study of these elements is biologically interesting as the mode of genetic propagation is kaleidoscopic and important, as MGEs are the main vehicles of the increasing bacterial antibiotic resistance that causes thousands of human deaths each year. The study of MGEs has previously focused on plasmids from individual isolates, but the revolution in sequencing technology has allowed the study of mobile genomic elements of entire communities using metagenomic approaches. The problem in using metagenomic sequencing for the study of MGEs is that plasmids and other mobile elements only comprise a small fraction of the total genetic content that are difficult to separate from chromosomal DNA based on sequence alone. The distinction between plasmid and chromosome is important as the mobility and regulation of genes largely depend on their genetic context. Several different approaches have been proposed that specifically enrich plasmid DNA from community samples. Here, we review recent approaches used to study entire plasmid pools from complex environments, and point out possible future developments for and pitfalls of these approaches. Further, we discuss the use of the PacBio long-read sequencing technology for MGE discovery.

Keywords: mobilome, plasmidome, PacBio, plasmid, horizontal gene transfer, metagenomics, mobile genetic elements, MGE

INTRODUCTION

Thanks to the enormous adaptability of pathogenic bacteria in clinical settings, public interest in microbiology has never been greater. Bacterial plasmids are the main basis of this adaptability and provide a platform for the recombination and transfer events that allow genes to spread horizontally between genera (Sørensen et al., 2005). This adaptability enables bacteria to acquire accessory genes such as antibiotic resistance genes and entire genetic pathways, e.g., encoding degradation of pesticides and other xenobiotic compounds in various environmental settings (Izmalkova et al., 2006; Dealtry et al., 2014).

Until recently, the study of plasmids in the environment was based on indicators of diversity using indirect and relative methods such as gradient gel resolution of PCR products, quantification of groups of incompatibility using qPCR, Southern blotting, or analysis of plasmids from cultivated transconjugants (Götz et al., 1996; Smalla et al., 2000). The recent revolution in sequencing technologies has provided researchers with a more nuanced view of mobile genetic elements (MGEs) in complex environments, since the cultivation and primer biases that previously limited research have been replaced by a different set of biases, yielding different insights. The sequencing based approaches aim to analyze entire MGE pools of communities or biological niches rather than select groups and are known by various names, such as mobilomics (Jørgensen et al., 2014), metamobilomics (Li et al.,

2012; Norman et al., 2014), plasmidomics (Kav et al., 2012) plasmid metagenomics (Palenik et al., 2009; Sentschilo et al., 2013), and viromics in the study of viruses (Lopez-Bueno et al., 2009). Though their applications differ slightly, these are, in essence, based on similar strategies, achieve the same type of results, and are restricted by some of the same biases.

Here, we discuss the biases and possibilities of current approaches and attempt to predict the direction in which this research is moving, specifically in regards to emerging sample preparation techniques, sequencing technology, and bioinformatics. The above methods are referred to collectively as mobilomics in this perspective article as not only gene-carrying plasmids, but also other circular elements such as phages and circular transposons are the target of these types of DNA purifications and as the increased mobility of these genetic elements relative to chromosomal DNA are pivotal for research interest in them.

MOBILOME SAMPLE METHODOLOGIES

In recent publications, there have been several approaches to obtain mobilome samples from complex environments, with the four most influential being A Cesium Chloride Ethidium Bromide (CsCl-EB) -gradient ultracentrifugation (Sentschilo et al., 2013), B bioinformatical derivation of plasmid related contigs/genes (Palenik et al., 2009; Ma et al., 2012), C transposon aided capture (TRACA; Jones and Marchesi, 2006), and D degradation of

linear DNA with exonuclease followed by multiple displacement amplification (MDA; Kav et al., 2012, 2013; Li et al., 2012; Jørgensen et al., 2014; Norman et al., 2014). In **Table 1**, the routes of the above studies are outlined, highlighting the similarity of approaches.

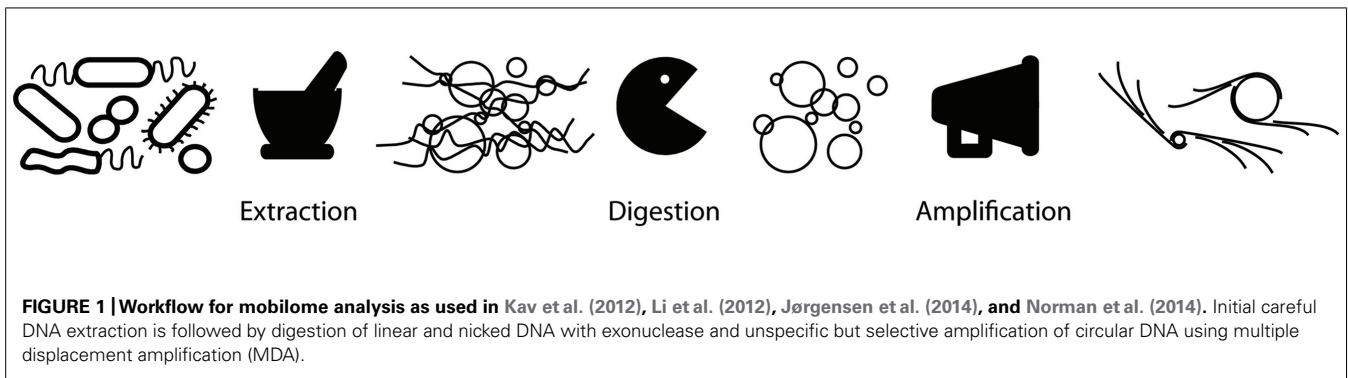
Approach A, CsCl-EB-gradient ultracentrifugation, is a classical and potent plasmid purification method, but the widespread use have diminished in the later decades after faster purification techniques became available in the early nineties. The method requires a large amount of starting material and specialized equipment, both of which are limiting factors for the use of this method (Sambrook et al., 1989). The advantage of this approach, though, is its potential to obtain large intact plasmids, a feat none of the other methods have been able to efficiently achieve. After careful DNA extraction and prolonged ultracentrifugation, supercoiled covalently closed circular DNA will stabilize and form a band in the centrifuge tube that can be harvested by needle puncture of the tube. This yields μg of plasmid DNA that can subsequently be sequenced. The accessory genes predicted on sequences resulting from CsCl-EB-gradient ultracentrifugation are largely unknown, pointing to the power of this method to capture a fraction of large plasmids that have not been studied before. In the only published study using this method (Sentchilo et al., 2013), a high fraction of 16S rDNA carrying chromosomal DNA was found in one of three samples, compared to the plasmid purities reported for the enzyme based method described below (approach D). Approach B, bioinformatical derivation of plasmid related genes/contigs, is regularly used and built into several bioinformatical pipelines

such as MG-RAST (Meyer et al., 2008). This metagenomic-based approach is commonly used when mobility is not the focus of study as only a small fraction of metagenomic datasets stem from MGEs. Further, it relies heavily on database adequacy and is thus most useful for studying well-characterized environments. Approach C, TRACA, rely on inserting a transposon carrying a selective marker on MGEs that are then used to transform bacteria. The transformants can then be grown selectively in culture and the MGEs extracted. Method C was promising when published, but has not gained popularity, possibly owing to its low throughput and inability to capture large plasmids outside of the range of the method described below. Further, it relies on the ability of *Escherichia coli* to maintain the plasmid thereby presumably excluding most plasmids from identification. For the above reasons and due to the fact that no published studies in the last 3 years have utilized this system, the method will not be touched upon further in this perspective article.

Approach D uses exonuclease to degrade presumably nicked chromosomal DNA and subsequently the rolling circle type MDA in order to enhance the signal from circular DNA and produce sufficient quantities of DNA for library construction (workflow illustrated in **Figure 1**). This approach has enabled analysis of complete circular sequences and produced numerous very pure mobilome samples. The mobilome sample purity has been estimated by qPCR on the 16s rRNA gene, where purities of upward of 99% are not uncommon in published and unpublished data (Kav et al., 2013; Jørgensen et al., 2014). After sequencing, mapping of reads on the Arb-Silva database of ribosomal DNA has

Table 1 | Routes for analyzing plasmid diversity in recent publications.

Study	Input material	Sample treatment	DNA purification	DNA treatment	DNA amplification	Sequencing platform	Confirmation of circularity
Jones and Marchesi (2006)	Human feces	Nycodenz	Alkaline lysis	Exonuclease Transposon aided capture (TRACA)	Transformation, culturing	Sanger, primer walking	PCR
Ma et al. (2012)	Seawater	Filtration, FACS	Phenol-chloroform	–	MDA	454	Inverse PCR
Kav et al. (2012)	Cow rumen	Homogenization, centrifugation, filtration	Alkaline lysis and lysozyme-SDS	Exonuclease	MDA	Illumina GAIIx	–
Li et al. (2012)	Wastewater	Sedimentation, centrifugation	Plasmid prep kit	Exonuclease	MDA	454	–
Sentchilo et al. (2013)	Sludge	Homogenization, centrifugation, pectinase	Alkaline lysis, acid phenol chloroform	CsCl-ethidium bromide density gradient ultracentrifugation	–	Sanger 454FLX	PCR
Jørgensen et al. (2014)	Rat cecum	RNAlater storage	Plasmid prep kit	Exonuclease	MDA	Illumina HiSeq2000 +454FLX	<i>In silico</i> and inverse PCR
Norman et al. (2014)	Wastewater	Filtration, centrifugation	Plasmid prep kit	Exonuclease, AGE size separation, electroelution	MDA	Illumina HiSeq2000	–



confirmed the accuracy of this method (<http://www.arb-silva.de/>, unpublished data). The estimation of chromosomal DNA contamination is particularly important as our extensive experience with building mobilome samples has led to acceptance of regular and seemingly stochastic failure in obtaining sufficiently low chromosomal contamination, often necessitating repetition of the entire procedure.

By analyzing DNA sequence data from mobilome samples, researchers have found a plethora of new, small plasmids and small plasmid fragments. Notably, though, very few plasmids larger than 10 kb have been completely assembled without using manual, PCR-based genome closing (Sentchilo et al., 2013; Jørgensen et al., 2014). Further, diversity of genes encoding plasmid stability (replication, mobilization, and toxin-antitoxin systems) outnumber diversity of accessory genes encoding, e.g., antibiotic resistance or substrate degradation, a striking difference from database plasmids that are often carrying multiple accessory genes (Kav et al., 2012; Jørgensen et al., 2014). Among the genes not related to essential plasmid functions, the proportion of genes encoding a protein with unknown function is high, often above 50% of predicted genes (Sentchilo et al., 2013). While the genetic composition of the found plasmids may, to some extent, reflect the reality in nature, it seems likely that known methodological biases contribute to this outcome. Thus, small plasmids are thought to be capable of carrying relatively few accessory genes and in every step of the predominant mobilome sample preparation (Figure 1), a new bias toward small plasmids is introduced (Jørgensen et al., 2014). One important such bias in mobilome sample construction is in the purification of DNA (Table 1, 'DNA treatment' and Figure 1 'Extraction' and 'digestion'). Because larger plasmids will be subjected to greater shear forces than smaller ones when in solution, they are more likely to break, and thus become substrate in the following exonuclease treatment. Further, different methods of extraction are thought to be more effective depending on particular bacterial groups, for example the Nycodenz gradient centrifugation method where both over- and underrepresentation of phyla have been observed (Holmsgaard et al., 2011). Another example is the plasmid prep extraction used in, e.g., (Li et al., 2012; Jørgensen et al., 2014). This plasmid extraction method is optimized for *E. coli* and is thought to underrepresent bacteria with a more rigid cell wall structure. This extraction bias have been attempted to be countered by pooling DNA purified with different methods (Kav et al., 2012). MDA bias are described in

detail below and presumed less important biases are described in more detail in a previous study from our group (Jørgensen et al., 2014).

The great advantage in using the enzyme based mobilome sample method depicted in Figure 1 is the purity of the resulting plasmid DNA, a purity that allows researchers to categorize all resulting reads and contigs as plasmid derived. This situation contrasts that of metagenomes, where reads are assumed to stem from chromosomes unless distinct plasmid-like features such as plasmid replication protein can be found, implying a bias toward previously characterized sequences. The mobilome sample method, on the other hand, is free of such biases, which allows exploration of entirely novel classes of circular MGEs. Moreover, as the polymerase used for MDA is able to amplify circular sequences multiple times for each binding forming a linear, direct-repeat structure, complete closed plasmids from a single plasmid can, theoretically, be identified. The closing of sequences into circular units are important as they represent complete biological units with full genetic context as opposed to bits and pieces such as reads, contigs, and scaffolds which might lack important contextual information (Jørgensen et al., 2014).

THE USE OF MULTIPLE DISPLACEMENT AMPLIFICATION IN PLASMID STUDIES

Multiple displacement amplification is widely used in single cell genomics and many other fields where the initial amount of DNA is insufficient for analysis. The method combines the tight binding and low error rate of the DNA polymerase from the phi29 phage with capped random primers causing strand displacement rather than digestion to produce up to 100 kb fragments with little sequence bias (Dean et al., 2001; Hutchison et al., 2005). The method has progressed from one-of-many amplification methods to a dominating position in unspecific amplification of DNA. The advantages are many: high yield, fast and easy preparation and the ability to re-amplify DNA with little to no sequence diversity loss, if the stock is running low (Dean et al., 2001; Hutchison et al., 2005). The amplification comes at a cost, however. In the case of plasmid amplification, the rolling circle mechanism allows the phi29 polymerase to selectively amplify circular molecules, but has long been hypothesized to skew abundance relation between individual circular molecules as each nucleotide on a short circular molecule is copied multiple times per polymerase binding, making quantitation efforts on MDA-based samples questionable at best.

A recent publication have confirmed and quantified this important method bias (Norman et al., 2014). While efficient in suppressing chromosomal ‘background’ DNA in samples, the small circular element bias will also impair the discovery of important larger plasmids carrying interesting accessory genes (Hutchison et al., 2005; Norman et al., 2014).

MOBILOME PREPARATIONS WITHOUT AMPLIFICATION

As MDA introduces a major bias in the mobilome sample preparation protocol (Norman et al., 2014), a natural consequence would be to omit the MDA step. Previously, this was not feasible because of input DNA quantity requirements to sequencing library construction, but with the emergence of enzyme-based DNA fragmentation, such as Nextera XT®, as little as 20pg DNA can be used as input, theoretically allowing the omission of MDA (Parkinson et al., 2012). The amount of DNA left after digestion and before MDA (**Figure 1**) is so minuscule that it is difficult to measure with standard methods such as Qubit HS fluorometry or spectrophotometry, and thus chromosomal contamination cannot be quantified prior to sequencing.

Although well documented, only one paper using the CsCl-EB-gradient method has documented success with library construction from unamplified mobilome samples to date. And unpublished results from our lab using the exonuclease-MDA method (**Figure 1**) suggest that the levels of chromosomal DNA in such samples are at least an order of magnitude higher than in MDA samples from the exact same DNA purifications, possibly owing to incomplete breakage/nicking of chromosomal DNA during purification. The high levels of chromosomal contamination in unamplified samples support the assumption that a major factor in achieving low chromosomal DNA content in mobilome samples is the MDA process itself. Contrary to expectations, sequencing of two unamplified mobilome samples has not yielded longer circular contigs, suggesting downstream pipeline shortcomings (unpublished data). These insufficiencies need to be investigated by more systematic testing and deeper sequencing. An alternative strategy to omitting MDA to include large plasmids in mobilome samples is size selection by gel electrophoresis and electro-elution (Norman et al., 2014). This approach aims to remove small plasmids before MDA, thus preventing them from claiming the lion’s share of rolling circle amplification. It is, however, yet to be explored fully with contemporary, sensitive, small volume methods, such as capillary electrophoresis, which minimizes cross-contamination and human error.

REPEATED STRUCTURES PROBLEM AND LONG READ SEQUENCING TECHNOLOGY

A recurring problem in the assembly of complete circular plasmids from complex communities is the many recombination events leading to repeated structures (Conlan et al., 2014). Thus, if an interspersed repeat is longer than the read length, most assemblers such as velvet (Zerbino and Birney, 2008) or IDBA-UD (Peng et al., 2012) will not be able to resolve the repeat beyond scaffolding, even with the help of paired end information (Jørgensen et al., 2014). This prevents automated recognition of many full circle plasmids, as it breaks large circular structures into several contigs. An example of this is seen in the only publication to

date that has succeeded in closing large plasmids from a complex environmental genetic pool (Sentchilo et al., 2013). Here, manual, PCR based genome closing was necessary to achieve the interesting results presented (Personal Communication). That PCR was needed for genome closing highlights both the ability of this method to successfully acquire information on large plasmids and the inability of assembly programs to automatically finish genomes.

Longer reads are seen by many as the solution to the repeat problems in plasmid sequencing. So far, only Pacific Biosciences (PacBio, Menlo Park, CA, USA) have made a commercial breakthrough with reads of up to 50 kb, though mostly much shorter in the range 1–10 kb. The technology has struggled with a poor quality of individual reads (low Phred score), but this could and should be corrected with either multiple readings of each fragment, high coverage or shorter, more reliable reads such as Illumina. Even with such correction, further problems with for PacBio plasmid sequencing are low output volume and the large amount of DNA required for sequencing (10–20 µg of DNA). Possibly, CsCl-EB-gradient ultracentrifugation could generate the required amounts of pure and unamplified plasmid DNA, but fruitful coupling of the two technologies remain to be shown.

A method to directly sequence plasmids on the PacBio platform with no library preparation and low quantity of input DNA has been suggested, but even in optimal conditions with a cultured mock community of two known plasmids has only been shown to yield 3.5% (479) plasmid derived reads from four PacBio RS runs (Coupland et al., 2012).

To test whether plasmid derived reads could be extracted from an MDA mobilome sample, we have submitted a well-defined sample (Jørgensen et al., 2014) to PacBio RS sequencing and analyzed the results by close manual dotplot inspection after error correction with Illumina reads and assembly with MIRA (Chevreux et al., 1999) (see supplementary materials and methods). The dot-plots seen in supplementary materials Figures S1A–C represent typical circular contigs, with A, representing a rare circular sequence of app. 2.5 kb and B, and C, representing the much more common multiple short repeats structure. From the assembly of one PacBio RS run yielding 24,711 reads with an N50 of 4392 nt, 12 entire plasmids smaller than eight kb could be identified whereas the rest of the potential circular contigs (214) consisted of numerous very short repeats (Supplementary Figures S1B,C), possibly representing random hexamer-based template independent products (Holbrook et al., 2005; Hutchison et al., 2005). Regardless of the origin, the contigs with extensive repeats and inverted repeats do not provide valuable information on gene-carrying plasmids in the environment. Since the sample is known to contain many hundreds of circular MGEs and only has approximately 1% chromosomal DNA, the low number of sequences that could be identified as circular points to severe problems in using MDA mobilome samples as input of PacBio sequencing.

The power of single molecule long read sequencing is indisputable and nicely illustrated by a recent paper identifying 21 large plasmids up to 320 kb in length from cultured, hospital-derived bacterial isolates (Conlan et al., 2014), using paired end information to close plasmids. Here, an impressive sequencing effort

proves that large plasmids, scarred by recombination events and with lots of repeated structures, can be completely assembled with technology available today. Critical here is the cultured nature of samples and the massive sequencing effort of four PacBio RSII runs per bacterial isolate in addition to Miseq and 454 sequencing and OpGen optical mapping (OpGen, Inc., Gaithersburg, MD, USA) to confirm results. Considering the substantial per base cost of PacBio sequencing and the difficulties in obtaining large quantities of pure plasmid DNA from environmental samples, the strategy is only feasible for cultivated organisms at the moment.

METAGENOMICS SPECIES APPROACH TO PLASMID ASSEMBLY

Very recently, it has been suggested that a co-abundance gene groups (CAGs) approach, where groups of genes are binned based on their abundance across multiple samples, could be used to identify plasmids (Nielsen et al., 2014). While this promising approach does not resolve the issue of repeated structures beyond scaffolding, it could allow for high quality scaffolds of plasmids to be analyzed. This could be a useful addition to the plasmid reference datasets liberating them from the current culturing bias, but at the same time introducing the uncertainty of not having intact, experimentally characterized sequences. Based on co-dependency of groups, valuable information on host relations of plasmids could be determined, although it is presumably most efficient for narrow host range plasmids. The weaknesses of the plasmid metagenomic species approach are the dependence on multiple samples and the enormous background of chromosomal DNA, necessitating very deep sequencing in complex samples. Further, while CAGs provide probability of contig coherence, it is not physical and irrefutable proof.

SUMMARY AND OUTLOOK

For reasons mentioned above, two main routes are realistic for mobilome studies in the coming years; if long read technology approaches the per-base cost of Illumina sequencing, it is possible that untreated metagenomes can be assembled into complete genomes and plasmids. Alternatively, researchers must fine-tune the sample preparation methods to avoid introducing the size bias of MDA to allow complete large plasmids to be analyzed, possibly by removing small plasmids from the pool prior to MDA. Either way, future bioinformatical pipelines must be designed for and constantly take into account the dynamic and circular nature of many bacterial extrachromosomal MGEs. A major factor in this would be for assembly programs to trivially resolve repeats where read pairs are bridging a number of contigs, effectively pushing the boundary of repeat disruption of complete plasmid assembly from the length of a read to the length that a read pair spans.

The study of mobilomes of complex samples has come a long way over the last few years, from PCR and culturing based approaches to analysis using high throughput sequencing and intricate bioinformatics pipelines to expose a world of plasmids and circular elements that previously went undetected. While far from exhaustive, these studies are paving the way for a more complete understanding of the diversity of plasmids and small circular elements, while also highlighting obstacles that are

to be overcome with elegant experimental and bioinformatical setups before the complete plasmid pools of environments can be explored and their significance for the survival and evolution of hosts and communities can be mapped.

ACKNOWLEDGMENTS

We thank Annelise Kjølner for comments and advise and the following institutions for financial support: The Lundbeck Foundation (www.lundbeckfoundation.com), grant project DK nr R44-A4384, the European Union's 7th Framework Program for Research and Technological Development (grant agreement no. 222625 METAEXPLORE, EU FP7 Theme KBBE-2007-3-3-05) and the Community Research and Development Information Service (CORDIS) The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <http://www.frontiersin.org/journal/10.3389/fmicb.2014.00750/abstract>

REFERENCES

- Chevreur, B., Wetter, T., and Suhai, S. (1999). "Genome sequence assembly using trace signals and additional sequence information," in *Proceedings of the German Conference on Bioinformatics*, Hannover, 99, 45–56.
- Conlan, S., Thomas, P. J., Deming, C., Park, M., Lau, A. F., Dekker, J. P., et al. (2014). Single-molecule sequencing to track plasmid diversity of hospital-associated carbapenemase-producing Enterobacteriaceae. *Sci. Transl. Med.* 6, 254ra126. doi: 10.1126/scitranslmed.3009845
- Coupland, P., Chandra, T., Quail, M., Reik, W., and Swerdlow, H. (2012). Direct sequencing of small genomes on the Pacific Biosciences RS without library preparation. *BioTechniques* 53, 365–372. doi: 10.2144/000113962
- Dealtry, S., Holmsgaard, P. N., Dunon, V., Jechalke, S., Ding, G. C., Krögerckenfort, E., et al. (2014). Shifts in abundance and diversity of mobile genetic elements to diverse pesticides introduced into an on-farm biopurification system over a year. *Appl. Environ. Microbiol.* 80, 4012–4020. doi: 10.1128/AEM.04016-13
- Dean, F. B., Nelson, J. R., Giesler, T. L., and Lasken, R. S. (2001). Rapid amplification of plasmid and phage DNA using phi29 DNA polymerase and multiply-primed rolling circle amplification. *Genome Res.* 11, 1095–1099. doi: 10.1101/gr.180501
- Götz, A., Pukall, R., Smit, E., Tietze, E., Prager, R., Tschäpe, H., et al. (1996). Detection and characterization of broad-host-range plasmids in environmental bacteria by PCR. *Appl. Environ. Microbiol.* 62, 2621–2628.
- Holbrook, J. F., Stabley, D., and Sol-Church, K. (2005). Exploring whole genome amplification as a DNA recovery tool for molecular genetic studies. *J. Biomol. Tech.* 16, 125.
- Holmsgaard, P. N., Norman, A., Hede, S. C., Poulsen, P. H., Al-Soud, W. A., Hansen, L. H., et al. (2011). Bias in bacterial diversity as a result of Nycodenz extraction from bulk soil. *Soil Biol. Biochem.* 43, 2152–2159. doi: 10.1016/j.soilbio.2011.06.019
- Hutchison, C. A., Smith, H. O., Pfannkoch, C., and Venter, J. C. (2005). Cell-free cloning using phi29 DNA polymerase. *Proc. Natl. Acad. Sci. U.S.A.* 102, 17332–17336. doi: 10.1073/pnas.0508809102
- Izmalkova, T. Y., Mavrodi, D. V., Sokolov, S. L., Kosheleva, I. A., Smalla, K., Thomas, C. M., et al. (2006). Molecular classification of IncP-9 naphthalene degradation plasmids. *Plasmid* 56, 1–10. doi: 10.1016/j.plasmid.2005.12.004
- Jones, B. V., and Marchesi, J. R. (2006). Transposon-aided capture (TRACA) of plasmids resident in the human gut mobile metagenome. *Nat. Methods* 4, 55–61. doi: 10.1038/nmeth964
- Jørgensen, T. S., Xu, Z., Hansen, M. A., Sørensen, S. J., and Hansen, L. H. (2014). Hundreds of Circular Novel Plasmids and DNA elements identified in a rat cecum metamobilome. *PLoS ONE* 9:e87924. doi: 10.1371/journal.pone.0087924

- Kav, A. B., Benhar, I., and Mizrahi, I. (2013). A method for purifying high quality and high yield plasmid DNA for metagenomic and deep sequencing approaches. *J. Microbiol. Methods* 95, 272–279. doi: 10.1016/j.mimet.2013.09.008
- Kav, A. B., Sasson, G., Jami, E., Doron-Faigenboim, A., Benhar, I., and Mizrahi, I. (2012). Insights into the bovine rumen plasmidome. *Proc. Natl. Acad. Sci. U.S.A.* 109, 5452–5457. doi: 10.1073/pnas.1116410109
- Li, L. L., Norman, A., Hansen, L. H., and Sørensen, S. J. (2012). Metamobilomics: expanding our knowledge on the pool of plasmid encoded traits in natural environments using high-throughput sequencing. *Clin. Microbiol. Infect.* 18, 5–7. doi: 10.1111/j.1469-0691.2012.03862.x
- Lopez-Bueno, A., Tamames, J., Velazquez, D., Moya, A., Quesada, A., and Alcamí, A. (2009). High diversity of the viral community from an Antarctic lake. *Science* 326, 858–861. doi: 10.1126/science.1179287
- Ma, Y., Paulsen, I. T., and Palenik, B. (2012). Analysis of two marine metagenomes reveals the diversity of plasmids in oceanic environments. *Environ. Microbiol.* 14, 453–466. doi: 10.1111/j.1462-2920.2011.02633.x
- Meyer, F., Paarmann, D., D'Souza, M., Olson, R., Glass, E. M., Kubal, M., et al. (2008). The metagenomics RAST server - a public resource for the automatic phylogenetic and functional analysis of metagenomes. *BMC Bioinformatics* 9:386. doi: 10.1186/1471-2105-9-386
- Nielsen, H. B., Almeida, M., Juncker, A. S., Rasmussen, S., Li, J., Sunagawa, S., et al. (2014). Identification and assembly of genomes and genetic elements in complex metagenomic samples without using reference genomes. *Nat. Biotechnol.* 32, 822–828. doi: 10.1038/nbt.2939
- Norman, A., Riber, L., Luo, W., Li, L. L., Hansen, L. H., and Sørensen, S. J. (2014). An improved method for including upper size range plasmids in metamobilomes. *PLoS ONE* 9:e104405. doi: 10.1371/journal.pone.0104405
- Palenik, B., Ren, Q., Tai, V., and Paulsen, I. T. (2009). Coastal *Synechococcus* metagenome reveals major roles for horizontal gene transfer and plasmids in population diversity. *Environ. Microbiol.* 11, 349–359. doi: 10.1111/j.1462-2920.2008.01772.x
- Parkinson, N. J., Maslau, S., Ferneyhough, B., Zhang, G., Gregory, L., Buck, D., et al. (2012). Preparation of high-quality next-generation sequencing libraries from picogram quantities of target DNA. *Genome Res.* 22, 125–133. doi: 10.1101/gr.124016.111
- Peng, Y., Leung, H., Yiu, S. M., and Chin, F. Y. L. (2012). IDBA-UD: a de novo assembler for single-cell and metagenomic sequencing data with highly uneven depth. *Bioinformatics* 28, 1420–1428. doi: 10.1093/bioinformatics/bts174
- Sambrook, J., Fritsch, E. F., and Maniatis, T. (1989). *Molecular Cloning*, New York: Cold spring harbor laboratory press.
- Sentchilo, V., Mayer, A. P., Guy, L., Miyazaki, R., Tringe, S. G., Barry, K., et al. (2013). Community-wide plasmid gene mobilization and selection. *ISME J.* 7, 1173–1186. doi: 10.1038/ismej.2013.13
- Smalla, K., Heuer, H., Götz, A., Niemeyer, D., Krögerrecklenfort, E., and Tietze, E. (2000). Exogenous isolation of antibiotic resistance plasmids from piggy manure slurries reveals a high prevalence and diversity of IncQ-like plasmids. *Appl. Environ. Microbiol.* 66, 4854–4862. doi: 10.1128/AEM.66.11.4854-4862.2000
- Sørensen, S. J., Bailey, M., Hansen, L. H., Kroer, N., and Wuertz, S. (2005). Studying plasmid horizontal transfer in situ: a critical review. *Nat. Rev. Microbiol.* 3, 700–710. doi: 10.1038/nrmicro1232
- Zerbino, D. R., and Birney, E. (2008). Velvet: algorithms for de novo short read assembly using de Bruijn graphs. *Genome Res.* 18, 821–829. doi: 10.1101/gr.074492.107

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Received: 31 October 2014; accepted: 10 December 2014; published online: 22 January 2015.

Citation: Jørgensen TS, Kiil AS, Hansen MA, Sørensen SJ and Hansen LH (2015) Current strategies for mobilome research. *Front. Microbiol.* 5:750. doi: 10.3389/fmicb.2014.00750

This article was submitted to *Evolutionary and Genomic Microbiology*, a section of the journal *Frontiers in Microbiology*.

Copyright © 2015 Jørgensen, Kiil, Hansen, Sørensen and Hansen. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.