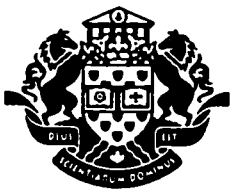


Université d'Ottawa · University of Ottawa

Université d'Ottawa · University of Ottawa

FACULTÉ DES ÉTUDES SUPÉRIEURES ET POSTDOCTORALES

FACULTY OF GRADUATE AND POSTDOCTORAL STUDIES

FILLION BERGERON, Marianne

AUTEUR DE LA THÈSE - AUTHOR OF THESIS

M.Sc. (Biology)

GRADE - DEGREE

Department of Biology

FACULTÉ, ÉCOLE, DÉPARTEMENT - FACULTY, SCHOOL, DEPARTMENT

TITRE DE LA THÈSE - TITLE OF THE THESIS

Plasmid Isolation and Purification by Electrofiltration and Comparison of Different Direct Colony Sequencing Methods and PCR-based Sequencing Methods

Guy Drouin

DIRECTEUR DE LA THÈSE - THESIS SUPERVISOR

EXAMINATEURS DE LA THÈSE - THESIS EXAMINERS

L. Bonen

D. A. Johnston

I.. Lambert

G. Slater

J.-M. De Koninck, Ph.D.

LE DOYEN DE LA FACULTÉ DES ÉTUDES SUPÉRIEURES ET POSTDOCTORALES

SIGNATURE

DEAN OF THE FACULTY OF GRADUATE AND POSTDOCTORAL STUDIES

Plasmid isolation and purification by electrofiltration and Comparison of different direct colony sequencing methods and PCR-based sequencing methods

Marianne Fillion Bergeron

Thesis submitted to the
Faculty of Graduate and Postdoctoral Studies
University of Ottawa
In partial fulfillment of the requirements for the
M.Sc degree in the

Ottawa-Carleton Institute of Biology

Thèse soumise à la
Faculté des études supérieures et postdoctorales
Université d'Ottawa
en vue de l'obtention de la maîtrise ès sciences

L'Institut de biologie d'Ottawa-Carleton

University of Ottawa
December 16th, 2001

ACKNOWLEDGMENTS

I realised as I went along my two years of master's that research is not as much an intellectual challenge as it is a mental one. As discouragement is often present, so are friends to keep you going. I would like to thank Dr. Guy Drouin for offering me the opportunity to work on the project, as well as my co-supervisor Dr. Gary Slater. A very warm thank you goes to all the labmates I have worked with over the years in this lab; thanks to Shirine, my partner in crime, even though you can't make decent coffee, Jamie, my molecular source of information, always calm and reassuring, Jenn, my moral support and yoga friend and Claude, my first lab trainer. I would also like to thank Dr. John Basso for all the encouragements, it was greatly appreciated. A very special thank you to my parents who have encourage me in my research, even though they could not understand what I was doing. A special hug for Van Gogh, my buddy, always there to distract me from my work. And to Marie, for listening over and over again about all my problems, work related or others. Finally, I would like to thank Applied Biosystems for providing funds for this project, as well as contributing to my OGSST scholarship.

TABLE OF CONTENTS

	Page
ACKNOWLEDGEMENTS	i
TABLE OF CONTENTS	ii
ABSTRACT	1
LIST OF FIGURES	3
LIST OF TABLES	4
Part I: Plasmid isolation and purification by electrofiltration	
CHAPTER 1. Introduction	8
1.1 General introduction	8
1.1.1 What is a plasmid	8
1.1.2 Uses of plasmids: molecular tools (mini-preps)	10
1.1.2.1 Cloning	10
1.1.2.2 Library screening	10
1.1.2.3 Sequencing	11
1.1.2.4 Gene therapy	11
1.2 Cell lysis	12
1.2.1 Alkaline lysis	12
1.2.2 Boiling lysis	13
1.3 Plasmid purification	14
1.3.1 CsCl ultracentrifugation	14
1.3.2 Chromatography	15
1.3.2.1 Ion-exchange chromatography	16
1.3.2.2 Magnetite	16
1.3.2.3 Triplex affinity interaction	17
1.3.3 96-well plates and automatization	18
1.3.4 Electrophoresis	19
1.3.4.1 Separation of plasmids	19
1.3.4.2 Electrophoresis of human genomic DNA	20

1.3.4.3 Electrofiltration of human genomic DNA 21
1.4 Hypothesis and Objective of mini-preps 22
1.4.1 Objective 22
1.4.2 Hypothesis 22

CHAPTER 2. Materials and Methods 23
2.1 Transformation 23
2.2 Boiling lysis 23
2.3 Electrofiltration 24
2.3.1 General electrofiltration conditions 24
2.3.2 Final electrofiltration conditions 24
2.4 Polymerase chain reaction 27
2.5 DNA purity and yield (spectrophotometer) 27
2.6 Calculation of the standard deviation of the samples (concentration and absorbance ratio) 28
2.7 Sequencing 29
2.8 Pure plasmid electrofiltration 29
2.9 Chromatographic purification 30
2.10 Proteinase K treatment 30

CHAPTER 3. Results 32
3.1 Boiling lysis buffer: STET at different concentrations 32
3.2 Conditions of electrofiltration 32
3.2.1 Time of electrofiltration 32
3.2.2 Buffer: TAE vs TBE 37
3.2.3 Concentration of buffer (TAE) 37
3.2.4 Membrane compositions and pore sizes 40
3.3 PCR dilution tests 44
3.4 Sequencing results of purified samples 45
3.5 Pure plasmid electrofiltration 48
3.6 Physical properties of electrofiltrated plasmid 49

3.7 Chromatography purification of electrophoresed samples 53
3.8 Proteinase K purification of electrophoresed samples 56

CHAPTER 4. Discussion 58
4.1 Electrofiltration: Effect of different parameters 58
4.2 PCR dilution test and sequencing 63
4.2.1 PCR dilution test: not informative 63
4.2.2 Sequencing: does not work with electrofiltration only 64
4.3 Purity of DNA purified by electrofiltration 65
4.3.1 STET, TAE (concentration) and boiling 65
4.3.2 Further purification of DNA isolated by electrofiltration 65
4.3.2.1 Chromatography purification 66
4.3.2.2 Proteinase K 67
4.4 Conclusion 68

Part II: Comparison of different direct colony sequencing methods and PCR-based sequencing methods

CHAPTER 5. Introduction 70
5.1 General introduction 70
5.1.1 Preparation of template for sequencing 70
5.2 Objectives of direct sequencing 70

CHAPTER 6. Materials and Methods 72
6.1 Direct colony sequencing I: Sasho (2000) 72
6.2 Direct colony sequencing II: Brown *et al.* (2000) 73
6.3 Direct colony sequencing III: Zhang *et al.* (1999) 73
6.4 PCR-based sequencing I: Silva *et al.* (2001) 74
6.5 PCR-based sequencing II: Smith *et al.* (2000) 75

6.6 PCR-based sequencing III: Truett *et al.* (2000) 76

CHAPTER 7. Results 79
7.1 Direct colony sequencing I 79
7.2 Direct colony sequencing II 79
7.3 Direct colony sequencing III 81
7.4 PCR-based sequencing I 81
7.5 PCR-based sequencing II 83
7.6 PCR-based sequencing III 83
7.7 Comparison of the six sequencing methods 87

CHAPTER 8. Discussion 89
8.1 Direct colony sequencing I, II and III 89
8.2 PCR-based sequencing I, II and III 90
8.3 Comparison of all six methods 91
8.4 Conclusion 92

REFERENCES 93

SUMMARY

We have designed an electrophoresis system that can purify plasmid DNA from a culture without centrifugation. This system is based on electrofiltration where bacterial cell lysates are loaded in one chamber and the purified plasmid DNA is recovered in an adjacent chamber. These two chambers are separated by a membrane made of regenerated cellulose, which allows plasmid DNA to migrate to the recovery chamber while retaining most contaminants in the loading chamber. Unfortunately, even with the optimization of the parameters involved in the electrofiltration, the only DNA that can pass through the middle membrane still has some contaminants, which prevent sequencing of the plasmid. Our results have shown that a pure plasmid cannot cross a membrane with pores small enough to prevent the migration of most of the contaminants. Only a plasmid complexed with some contaminants can cross a small pore membrane. In parallel, we have compared six direct sequencing methods that do not require any plasmid purification prior to the sequencing reaction. We compared the reliability, quality of sequences, time required, and cost of these six methods. We found that the best method was that of Zhang *et al.* (1999). This method is fast, reliable, produces good quality sequences and is inexpensive. The performance of this method is due to the amount of ABI's ready reaction mix used, the pre-sequencing heating step to lyse the cell, the large volume of the PCR sequencing reaction and the addition of BSA.

RÉSUMÉ

Nous avons construit un système d'électrophorèse pour purifier l'ADN plasmidique d'une culture de bactéries, sans centrifugation. Ce système est basé sur l'électrofiltration, où un lysat de bactéries est chargé dans une chambre de départ; l'ADN plasmidique purifié est récupéré dans la chambre d'arrivée. Ces deux chambres sont séparées par une membrane faite de cellulose régénérée, qui permet à l'ADN plasmidique de migrer vers la chambre d'arrivée, tandis que les contaminants sont retenus presqu'en totalité dans la chambre de départ. Malheureusement, même lorsque tous les paramètres impliqués dans l'électrofiltration sont optimisés, seul l'ADN plasmidique contenant certains contaminants peut traverser la membrane, ce qui rend le séquençage du plasmide impossible. Nos résultats ont démontré qu'un plasmide pur ne peut pas traverser une membrane avec des pores assez petits pour prévenir la migration de la plupart des contaminants. Le seul plasmide qui peut traverser une membrane de petits pores est un plasmide qui est associé avec des contaminants. Cependant, nous avons effectué une comparaison de six méthodes de séquençage directes qui ne requièrent pas de purification de plasmide avant la réaction de séquençage. Nous avons comparé la fiabilité, la qualité des séquences, le temps de travail nécessaire, ainsi que le coût des six méthodes. Nos résultats ont montré que la meilleure méthode est celle de Zhang *et al.* (1999). Cette méthode est rapide, fiable, produit des séquences de qualité et est peu coûteuse. La performance de cette méthode est due à la quantité de «ABI ready reaction mix » utilisée, à l'étape de chauffage avant le séquençage pour lyser les cellules, le volume important de la réaction de polymérase en chaîne et l'ajout de BSA.

LIST OF FIGURES

	Page
Figure 1. Electrophoresis tank	25
Figure 2. Schematic of the chamber system.	26
Figure 3. Time of electrofiltration	34
Figure 4. Backpulse time	36
Figure 5. PCR dilution test	47
Figure 6. Plasmid supercoiling after electrofiltration	54
Figure 7. Schematic of a polycarbonate and cellulose membrane	61
Figure 8. Silva *et al.* (2001) PCR comparison	82
Figure 9. Smith *et al.* (2000) PCR comparison	84
Figure 10. Truett *et al.* (2000) PCR comparison	85
Figure 11. Truett *et al.* (2000) alkaline lysis	86

LIST OF TABLES

Page

Table 1. Electrofiltrations with a 14kD ×2 RC middle membrane for 10, 20, 30 and 40 minutes. 35

Table 2. Average concentration (μg/μL) and A260/A280 ratio of recovered samples after electrofiltrations with two different buffers, 0.5 × TAE and 0.5 × TBE and a permissive middle membrane (300kD PC). 38

Table 3. Average conentration (μg/μL) and A260/A280 ratio of the recovered sample after electrofiltrations were performed with a 25kD CA middle membrane at 0.1 ×, 0.2 ×, 0.3 ×, 0.5 × and 1.0 × TAE. 39

Table 4. Average concentration (μg/μL) and ratio of the recovered sample after electrofiltrations performed with 2 different 300kD middle membrane, polycarbonate and cellulose acetate. 41

Table 5. Comparison of different middle membranes for electrofiltration according to the average concentration (μg/μL) and ratio of recovered samples. 43

Table 6. Comparison of different combinations of middle membranes of small and big pores with the average concentration (μg/μL) and ratio of the recovered samples. 46

Table 7. List of various controls for sequencing reactions with pure plasmid. 50

Table 8. Average concentration (μg/μL) and ratio of pure plasmid from the loading and recovery chambers after electrofiltration with different membranes. 51

Table 9. Average concentration (μg/μL) and ratio of pure plasmid from the loading and recovery chambers after electrofiltration of 10, 20 and 30 minutes. 52

Table 10. Chromatography purification of electrofiltrated samples. 55

Table 11. Concentration (μg/μL) and ratio of recovered samples with a 14kD ×2 RC or a 50kD PC membrane before and after being treated with proteinase K. 57

Table 12. Comparison between the modified PCR protocol and the three PCR protocols from the PCR-based methods. 78

Table 13. Success rate, quality, time and cost comparison of six novel sequencing methods. 80

LIST OF ABBREVIATIONS

BSA: bovine serum albumine

C: PCR negative control

CA: cellulose acetate

L: loading chamber

MW: molecular weight marker

PC: polycarbonate

PCR: polymerase chain reaction

PP: pure plasmid

R: recovery chamber

RC: regenerated cellulose

V: pure vector

Part I: Plasmid isolation and purification by electrofiltration

CHAPTER 1. Introduction

1.1 General introduction

In these glory days of molecular biology, most organisms' genomes are looked at for all kinds of purposes, whether they are medical, agricultural, evolutive or just for the quest of knowledge. These organisms have genomes of millions to billions of bases, and these millions of bases need to be broken down to obtain manageable research sizes. The recombinant DNA technology has allowed the studies of only specific genes or regions at a time, amongst the vast pool of an organism's sequences.

1.1.1 What is a plasmid?

This recombinant DNA technology has become possible due to the existence of plasmids, which are extrachromosomal DNA molecules that replicate independently inside a host cell. Plasmids are naturally found in some bacteria, and their size can vary from a couple of thousand bases to more than ten thousand bases (10kb) (Lodish *et al.* 1996). A plasmid replicates along with the host cell and can therefore give copies of itself to both daughter cells, propagating identical copies of the original plasmid through the bacterial lineage. Some plasmids also have the necessary genes to perform gene transfer, i.e. to replicate their DNA and to pass it on from one host cell to another (Lodish *et al.* 1996). Most plasmids used in laboratories have been extensively modified to accommodate all the needs required for various uses, such as reduction of the plasmid to keep only the essential genes and to make it easier to handle, and the addition of selective genes that simplify selection and procedures in the laboratory (Lodish *et al.* 1996). These

modified plasmids are used as vectors to isolate, amplify and move around DNA fragments by ligation of these fragments in the vector. As the plasmid replicates itself, the foreign DNA is replicated as well.

Plasmids are mostly used in two forms, i.e. supercoiled (in their native form) and linearized (e.g., by enzymes). Another form is the open circle, which is between the supercoiled and linear forms. By introducing a nick in one of the strands of the supercoiled double-stranded (ds) plasmid, the structure is relaxed and it forms an open circle. When adding a second nick at the same position to the double-stranded plasmid, it goes from the open circle to the linear form. In bacteria, most plasmids are supercoiled because this form is the most compact. The more supercoiling there is in a plasmid, the more compact it is. Most circular DNA molecules within cells are negatively supercoiled (Prazeres *et al.* 1998), but the positively supercoiled conformation is also possible. For a given molecular weight, the supercoiled conformation has the highest velocity (in comparison to an open circle or a linear molecule) when electrophoresed in an agarose or polyacrylamide gel. The open circular conformation of plasmids has not been studied as much as the supercoiled or the linear forms, probably due to its fewer uses in molecular biology. Finally, even though plasmids are not by definition linear, but rather closed circles, they are often linearized for multiple purposes such as molecular markers, cloning and sequencing. Obviously, the linear plasmids migrate in a gel the same way DNA fragments do. Their velocity is slower than the supercoiled form for a given molecular weight, but faster that the relaxed form because they are less easily trapped (impaled) by dangling gel fibers (agarose and polyacrylamide).

1.1.2 Uses of plasmids: molecular tools (mini-preps)

Plasmids are widely used in molecular biology as tools to perform a number of tasks, but mainly to amplify and move around pieces of foreign DNA (DNA that does not come from the plasmid itself). These applications include cloning, library screening, sequencing, and more recently, gene therapy.

1.1.2.1 Cloning

In order to isolate DNA fragments from a complex mix of fragments, the fragments need to be cloned into plasmid vectors, and these vectors are then transformed into a bacterial host (e.g., *E. coli*). When the bacterial host replicates on a plate or in a liquid culture, the plasmid containing the DNA fragment is also replicated, and a sufficient amount of DNA is produced and extracted from the bacterial cell to perform further studies on that particular sequence (Maniatis *et al.* 1989).

1.1.2.2 Library screening

Plasmids are also widely used for the purpose of library screening. When searching for a particular gene or piece of DNA in a whole genomic DNA fraction, it is useful to cut the genomic DNA into small fragments and to ligate all the pieces generated into different plasmids. Once the clones are done, they are transformed into bacterial host cells so the plasmid can replicate inside the host cells and amplify the cloned DNA along with its own DNA. This process gives rise to a library of vectors containing all the different pieces of the original genomic DNA, and the colonies can then be picked and grown to be screened.

1.1.2.3 Sequencing

To purify DNA for sequencing, the DNA first needs to be cloned into a plasmid vector, which is grown in a host cell, after which the plasmid is purified and used as the sequencing template. Moreover, if a DNA fragment from which no sequences are known needs to be sequenced, primers designed from the plasmid sequence itself and flanking the cloning region can be used as a start position for sequencing.

1.1.2.4 Gene therapy

As more and more genes are identified as being part of the cause for genetic diseases, more and more efforts are directed at the possibility of replacing, or compensating for, the defective genes. As a result, a new focus has been made on plasmids as tools or delivery vehicles for correcting genes inside human cells, and, following this, a new look has been posed on purification of plasmid DNA for pharmaceutical purposes (Levy *et al.* 2000). The protocols that are known and used now have been mostly developed for laboratory purposes only, and most of them cannot be applied for use in gene therapy. Although plasmid DNA seems to have great potential for both vaccines and medical therapy, some practical issues still exist regarding the isolation and purification protocol. New protocols specific for gene therapy need to be developed that will minimize levels of contaminants such as endotoxins, RNA, protein and bacterial host DNA to nothing (Ferreira *et al.* 2000, Prazeres *et al.* 1999). Moreover, since most plasmid purification protocols call for the use of toxic chemicals, the residual presence of these substances in the final plasmid solution will also need to be assessed. Finally, in order to make these new therapeutics available to everyone, the new protocols will need

to be very efficient regarding the yields of purified plasmid (Ferreira *et al.* 2000, Levy *et al.* 2000, Prazeres *et al.* 1999).

1.2 Cell lysis

After a bacterial host cell has been transformed with a plasmid and used as an incubator to amplify the plasmid, it is necessary to extract the latter from the host cell to be able to use it for further analysis or experiments. The first step for the purification of plasmids (also commonly called mini-preps) is the lysis of the bacterial cell, to release the plasmid from the host (Maniatis *et al.* 1989). In order to do that, the cell wall needs to be weakened or broken and, along with the plasmid, the whole cell contents are released. For more then twenty years now, two protocols have been used to break the cell wall, the alkaline lysis and the boiling lysis methods (Maniatis *et al.* 1989).

1.2.1 Alkaline lysis

The first protocol that was specifically described to release plasmid from bacterial cells, and to separate the plasmid from the other cell components, was the alkaline lysis method of Birnboim and Doly in 1979. After the bacterial cells have been transformed with the vectors containing fragments of foreign DNA, the colonies are grown on selective plates and the desired colony is picked and used to inoculate Luria Broth (LB) media. After an overnight growth of the culture at 37°C, the culture is centrifuged to form a solid bacterial cell pellet, and the supernatant is removed. The lysis of the bacterial pellet is composed of three major steps: resuspension, lysis and precipitation. The resuspension consists of resuspending the bacterial pellet in the appropriate buffer

(glucose, Tris-HCl, EDTA) (Birnboim and Doly, 1979). After the tube content are thoroughly mixed, the bacterial cells are lysed with the addition of a solution of NaOH and an ionic detergent (SDS) that not only breaks the cell wall, but also denatures the double-stranded DNA (bacterial genomic and plasmid) and the proteins. SDS solubilizes the phospholipid and protein components of the cell membrane, leading to lysis and release of the cell contents while the alkaline conditions denature the chromosomal and plasmid DNAs, as well as the proteins (Qiagen, 1998). Finally, to separate the plasmid from the rest of the bacterial components, a high salt buffer (containing potassium acetate and glacial acetic acid) is added. This buffer causes the precipitation of the denatured proteins along with the chromosomal DNA, cellular debris and SDS (Qiagen, 1998). The denatured plasmids are much smaller in size than the chromosomal DNA and under high salt conditions, renature and stay in the supernatant. Depending on what the plasmids have been isolated for, further purification can be done, e.g., using a phenol:chloroform step as mentioned in Birnboim and Doly (1979). The alkaline lysis method is used in most mini-prep protocols, and is usually followed by a diversity of further purification steps.

1.2.2 Boiling lysis

The boiling lysis method was published in 1981 by Holmes and Quigley. This method requires the same first steps as the alkaline method, i.e., the inoculation of LB media, overnight growth and centrifugation to remove the supernatant. The lysis also includes three major steps, which are the resuspension of the bacterial pellet, the addition of lysozyme and the boiling. The bacterial pellet is first resuspended in a STET solution

(NaCl, Tris-HCl, EDTA and Triton X-100) that acts as a buffer and a lysing agent; the Triton X-100 is a non-ionic detergent that weakens the bacterial cell wall. The second step is the addition of lysozyme, an enzyme that lyses the bacterial cell wall. Finally, the bacterial pellet is boiled for 40 seconds (Holmes and Quigley, 1981). After a low speed centrifugation, the plasmid remains in the supernatant while the insoluble clot of genomic DNA and debris forms a pellet that can be removed and the plasmid is recovered from the supernatant by isopropanol:ethanol precipitation (Holmes and Quigley, 1981).

1.3 Plasmid purification

After plasmids are isolated from most of the contaminants, it can be necessary to purify them more depending on what they will be used for (Maniatis *et al.* 1989). Of the two lysis methods, alkaline lysis has been adopted by most protocols and commercially available kits (Itoh *et al.* 1997, Prazeres *et al.* 1998, Qiagen 1998). Although the primary steps of alkaline plasmid isolation and purification (commonly called mini-preps) have not been modified much in the last twenty years, the protocols used to obtain pure supercoiled plasmid have evolved extensively. The following sections present some of these methods.

1.3.1 CsCl ultracentrifugation

One of the first methods that was designed to obtain ultrapure supercoiled DNA is the CsCl/EtBr ultracentrifugation. After alkaline lysis, supercoiled plasmid DNA is separated from the other contaminants such as linear DNA, proteins and RNA because different amounts of EtBr intercalate in the different forms of DNA. The sample

containing the supercoiled plasmid DNA and the contaminants is centrifuged at high-speed (from 45 000 rpm to 60 000 rpm depending on the centrifuge model) for periods ranging from twelve hours to two days (Maniatis *et al.* 1989). The centrifugation allows the separation of the different components of the loaded sample at various places in the centrifuge tube. Because EtBr is lighter than DNA density, the more molecules that are intercalated in the double helix, the lighter is the DNA. The supercoiling of the plasmid DNA makes it harder to incorporate ethidium bromide molecules into the double helix, because with every additional EtBr molecule, the supercoiling becomes tighter. Therefore, a supercoiled plasmid will have less EtBr incorporated than a linear plasmid, and it will have a higher density, forming a distinctive band in the tube lower than the linear DNA. However precise this method is, it is still very expensive, time-consuming, and it involves working with a fair amount of a carcinogenic agent (EtBr) (Maniatis et al, 1989). This method is therefore seldom currently used, especially since the availability of reliable chromatographic protocols.

1.3.2 Chromatography

Chromatography is currently one of the most popular methods for the purification of plasmids. It is also the method used in almost all of the commercially available kits. It can be modified for various amounts of plasmid recovery, going from 1mL of culture to 100 mL, and it can be used to purify plasmids of various sizes, from 3kb to megabase vectors. It can be adapted to large-scale production, whether it is in number of samples or in quantities produced. Various forms of chromatography exist, but they are all based on the same principle: under specific conditions, double-stranded plasmid DNA will bind to

a solid support, and following a shift in these conditions, the plasmid DNA will be released from the support and eluted.

1.3.2.1 Ion-exchange chromatography

Ion-exchange chromatography (also called anion-exchange) is based on the interaction between the negatively charged plasmid DNA and a positively charged matrix. After the last step of alkaline lysis and the following centrifugation, plasmid DNA is in a high salt solution while the other contaminants have been removed by centrifugation. Under high salt conditions, the plasmid DNA is selectively adsorbed unto a positively charged ligand (a solid support matrix) while the other contaminants remaining in the solution can be eluted with a centrifugation step. Once the remaining contaminants have been discarded, the high salt solution is replaced by a low salt solution, and the supercoiled plasmid is released from the matrix (Prazeres *et al.* 1998, Qiagen 1998). Most commercially available mini-prep kits are using this method with various matrices.

1.3.2.2 Magnetite

Other solid supports can be used in ion-exchange chromatography that do not require any centrifugation, which is usually the most time-consuming step when processing a large number of samples. Magnetite has been used as a solid-phase adsorbent in ion-exchange chromatography, and it works on the same principle as does a non-magnetizable matrix: double-stranded DNA is adsorbed on the magnetite support (Fe_3O_4) at high salt concentration, and it can be water eluted (low salt concentration)

once the contaminants have been washed away (Davies *et al.* 1998). Since the magnetite support (and the attached DNA) can be retained in the tube by a magnet, the high salt solution can simply be poured out from the tube, removing the need for centrifugation, and the only DNA remaining in the tube is the plasmid DNA (Skowronski *et al.* 1999). The same procedure is done for the water elution step, which makes this protocol very inexpensive if the magnetite support is made in the laboratory (Davies *et al.* 1998).

1.3.2.3 Triplex affinity interaction

The ion-exchange chromatographic columns discussed in the previous section are aimed at processing as many samples as possible in the shortest time possible, while some methods which are more gene therapy oriented at processing a large amount of one particular sample, and at reducing the need for chemical products. The triplex affinity chromatographic column is one of the novel methods that has been designed to purify one particular plasmid in large amounts, and it can be scaled-up easily. This method works on the principle of triplex affinity interaction, which means the formation of a triplex-DNA helix (Schluep and Cooney, 1998). The solid support is made of Sephacryl S-1000 beads to which a 15mer DNA target sequence is attached. This 15 base oligo-nucleotide is a single-stranded homopyrimidine strand that is designed to form a triple helix with a double-stranded homopurine-homopyrimidine complementary sequence. The third homopyrimidine strand binds to the major groove of the double helix, parallel to the homopurine strand of the Watson-Crick double helix with Hoogsteen hydrogen bonds (Schluep and Cooney, 1998). For gene therapy purposes, the fifteen base pairs homopurine sequence is cloned into the plasmid to be purified, and that plasmid, along

with the impurities, is mixed with the Sephacryl beads. Under acidic and high salt conditions, the triple-helix bonds are formed, and the plasmid is attached to the support. The impurities can be washed out and the pH is changed to alkaline conditions to release the plasmid from the ligand (Schluep and Cooney, 1998). This system can be used as a continuous affinity purification process, reusing the same column to purify large quantities of a specific plasmid.

1.3.3 96-well plates and automatization

With the advancement in sequencing technologies and the numerous genome sequencing projects, the need for high-throughput plasmid purification has been increasing over the last few years. Most cases of mini-preps implicating high number of samples are for sequencing purposes, and the required plasmid quality is high, leading to the use of chromatographic columns. These methods have been adapted to 96-well plate systems (Itoh *et al.* 1997, Itoh *et al.* 1999), and in some cases, the multiple time-consuming centrifugations have been replaced by vacuuming processes, which are faster, require less hands-on work and are easier to automate (McNally, 1999). This replacement of centrifugation along with the use of 96-well plates has allowed companies to come up with fully automatized setups that can process a very large number of samples in a very short time with almost no manipulation. However, these new technologies are not very wide-spread in common laboratories, mostly due to the high cost of the robots and materials.

1.3.4 Electrophoresis

Electrophoresis is mostly used in molecular biology as a mean to separate and to visualize DNA fragments of different sizes (down to 1 base difference). All DNA, whether it is single-stranded, double-stranded, supercoiled or linear, has a total negative charge. This property has been used to separate DNA in sieving matrices (e.g., agarose and polyacrylamide gels) by applying a voltage that makes DNA fragments migrate through the gel at a velocity proportional to their sizes. Electrophoresis has been used to purify plasmid DNA as well as genomic DNA (Cao and Ing 1991, Cole 1998, Cole 1999).

1.3.4.1 Separation of plasmids

In some gene therapy plasmid systems, the presence of a second plasmid in the bacterial host is necessary to allow the proper replication of the plasmid of interest (i.e. the plasmid carrying the replacement or compensatory gene). For example, the supF-containing plasmid requires the presence of the helper plasmid p3 to grow in the bacterial host (Prieto *et al.* 2000). Standard mini-prep procedures are not designed to differentiate between two plasmids during purification, and the two plasmids are then co-purified, which is undesirable in gene therapy. A new electrophoresis-based method can allow the purification of the desired plasmid, and allow it to be separated from the helper plasmid at the same time. An agarose gel is cast in a plastic tube with a 3-5mm long 2% agarose base layer and a 8cm long 1% agarose core. The plasmid solution is electrophoresed overnight through the agarose tube, eluted, and run through a UV detector and a fraction collector. The 2% agarose base layer serves to compact the DNA collected and

concentrate the samples (Prieto *et al.* 2000). Since the two plasmids do not have the same size, the agarose tube will separate the plasmids, and they will elute at different times. This process has the advantages of being affordable and easy, but it requires an overnight electrophoresis.

1.3.4.2 Electrophoresis of human genomic DNA

Electrophoresis can also be used to purify human genomic DNA from blood, and a 96-well system has been developed in our laboratory (Slater *et al.* 2000). For human genomic DNA purification, electrophoresis is used to separate the DNA from the contaminants instead of centrifugation. In a vertical agarose gel system that has wells on top of the gel, lysed blood samples are loaded in the wells and the electrophoresis is carried out according to a specific ratio of voltage and time in both directions. The goal is to make the DNA and the contaminants migrate into the gel during the forward (F) pulse. During the backward (B) electrophoresis, the genomic DNA goes back in the wells while the contaminants remain in the gel. When $V_F T_F = V_B T_B$ with $V_F/V_B = 2.4$ (where V_F: voltage forward, T_F: duration of forward pulse, V_B: voltage backward, T_B: duration of backward pulse), the genomic DNA is separated from the contaminants which are entrapped in the agarose gel, while the genomic DNA stays in the wells. Since the human genomic DNA is very long compared to the contaminants, under the forward pulse, it gets instantly trapped on the surface of the well while the contaminants migrate into the gel. Under the backward pulse, the genomic DNA is released from the well surface and returns in the loading well. As for the contaminants, they do migrate back up towards the loading well, but because of the precise field conditions chosen, they stay trapped in the

gel. This sequence is repeated for a number of cycles until all the contaminants have migrated deep inside the agarose gel.

1.3.4.3 Electrofiltration of human genomic DNA

Since the use of gels to purify human genomic DNA is not easily commercialized, electrophoresis was also used with another separating medium. Membranes and teflon screws were used to create chambers, a loading chamber where the blood lysate is loaded, and a recovery chamber where the human genomic DNA is recovered. When a voltage is applied, the negatively charge genomic DNA crosses the middle membrane separating the two chambers, while the contaminants are blocked by the same membrane, and stay in the loading chamber (Zhou *et al.*, unpublished). This system led us to try the same approach for plasmid purification, as described below.

1.4 Hypothesis and Objectives of the mini-prep project

1.4.1 Objective

The objectives of the first part of this work are to design a protocol to isolate and purify plasmids from cultures that does not require centrifugation, that is time-efficient, automation friendly, and that can produce a plasmid that will meet the requirements for sequencing (i.e., high sample concentration and purity).

1.4.2 Hypothesis

Based on previous results on the electrofiltration of human genomic DNA (Hongyan Zhou, Gary W. Slater, Guy Drouin, unpublished), we hypothesize that plasmids can also be separated from contaminants using membranes and electrophoresis.

CHAPTER 2. Materials and Methods

2.1 Transformation

The DH5α *E.coli* cells were from the TOPO F' cloning kit (Invitrogen) and were transformed according to the manufacturer's instructions with the plasmid vector pBluescript SK+II (0.376 μg/μL) (Invitrogen) that had been previously purified by CsCl/EtBr ultracentrifugation (Dr. Guy Drouin). The plasmid contained no insert and was 2961 bp long. The transformed cells were plated on LB agar plate containing ampicillin (5 μg/mL), X-gal (0.04 mg/mL) and IPTG (0.0005 M). The plates were incubated at 37°C overnight and blue colonies were picked from the plate to be grown in LB liquid media containing 5 μg/mL of ampicillin. The cultures (5 mL) were incubated overnight at 37°C and grown to an OD (A_{600}) (Pharmacia Biotech Novaspec II) of 0.400 to 0.700 (reference: Luria-Broth media).

2.2 Boiling lysis protocol

The boiling lysis was the cell lysis protocol chosen because it does not require centrifugation and because of the low salt concentration involved in it (high salt concentration can interfere with the electrofiltration and break the membranes). 110 μL of overnight culture (not centrifuged) was mixed with 35 μL of STET 8 (0.8 M NaCl, 80 mM Tris-HCl (pH 8.0), 8 mM EDTA (pH 8.0), 40% v/v Triton X-100), 4 μL of lysozyme-4 (40 mg/mL in 10 mM Tris-HCl pH 8.0) and 1 μL of RNase A (10 mg/mL in water; Boehringer Mannheim) and boiled (100°C) for 5 minutes.

2.3 Electrofiltration

2.3.1 General electrofiltration conditions

Unless stated otherwise, electrofiltrations were carried out at 250V for 10 to 60 minutes in 0.5× TAE with a final backpulse of 30 seconds at 250V in an electrofiltration tank (Figure 1). The end membranes were hand cut in 6-8kD regenerated cellulose (RC) (Spectrum) membranes. The pore size of the membranes are given in KiloDaltons, which is based on the size of the proteins that can cross the membrane through the pores. The middle membrane was different depending on the experiment: regenerated cellulose membranes (Spectrum) were hand cut, whereas polycarbonate (PC) and cellulose acetate (CA) were pre-cut membranes from Amika Corp. Four teflon screws placed in series (Amika Corp.) form the loading and recovery chambers that are placed in the electro-elutor buffer tank (Amika). The chambers have a volume of 130 μL; the loading chamber is filled with cell lysate, and the recovery chamber is filled with electrophoresis buffer (Figure 2).

2.3.2 Final electrofiltration conditions

Electrofiltrations to obtain sequencing grade plasmid were carried at 250V for 10 minutes with a backpulse of 30 s at 250V. The end membranes were 6-8kD regenerated cellulose (Spectrum) and the electrophoresis buffer was 0.5× TAE while the middle membrane was a double layer of 14kD regenerated cellulose (Spectrum).

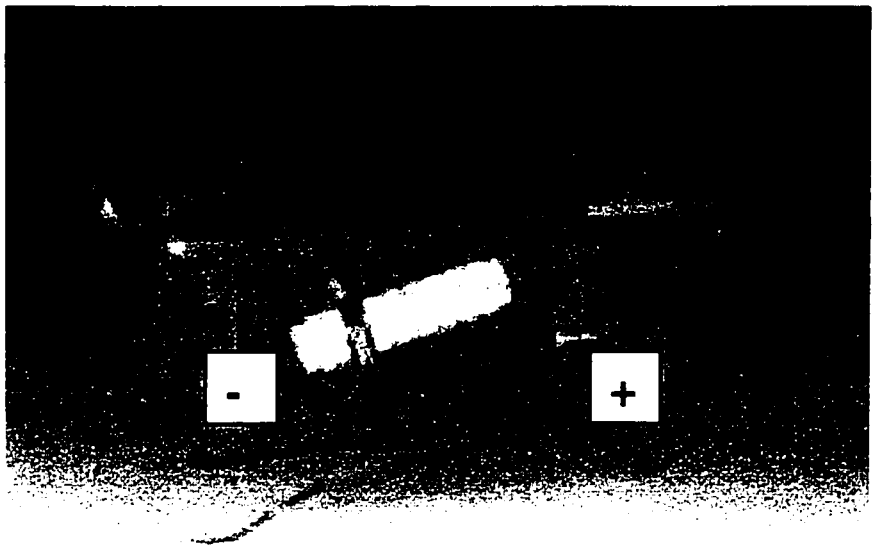


Figure 1. The electrophoresis tank in which the electrofiltrations were performed (Amika Corp).

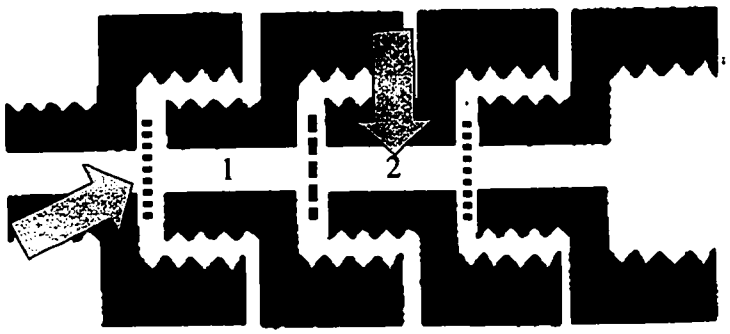


Figure 2. Schematic transversal view of the four teflon screws and the three membranes that define the loading (1) and recovery (2) chambers.

2.4 Polymerase Chain Reaction

The PCR protocol was the same for all PCR experiments done in the first part of the thesis. Five µL of DNA were mixed to 5 µL of dNTPs (5 mM) (Boehringer Mannheim), 5 µL of M13 forward primer (5 µM; Life Technologies), 5 µL of M13 reverse primer (5 µM; Life Technologies), 5 µL of 10× PCR buffer (Amersham Pharmacia Biotech Inc.), 2 µL of $MgCl_2$ (25 mM) (Boehringer Mannheim), 0.5 µL of Taq DNA polymerase (50 000 U/mL, Amersham Pharmacia Biotech Inc.) in a total volume of 50 µL (adjusted with deionazed water). The PCR is done on a Mandel Scientific Progene thermal cycler at 94°C for 3 min, then for 20 to 30 cycles at 94°C for 1 min, 46°C for 1 min and 72°C for 1 min, followed by 1 cycle at 72°C for 3 min. The PCR products were visualized by running them on a 1% agarose gel.

2.5 DNA purity and yield (spectrophotometer)

A Pharmacia Gene Quant spectrophotometer was used to measure the DNA absorbance ratio (A_{260}/A_{280}) and concentration (µg/µL). The absorbance ratio that we want to obtain needs to be as close to 1.800 as possible, which indicates pure DNA; if the number is different, it indicates the presence of contaminants in the samples. For electrofiltrated plasmids, the reference used to measure absorbance ratio was: 110 µL of 0.5× TAE, 35 µL of STET 8 electrofiltrated under the same conditions i.e., using the same time and middle membrane as the electrofiltrated plasmid.

2.6 Calculation of the standard deviation for our data (concentration and absorbance ratio)

In light of the high variability of our results, a standard deviation calculation was used to better characterize our results. These calculations were a means to assess whether the difference observed between the mean values obtained for the different experimental conditions was meaningful when compared to the scattering in the data points. The standard deviation s is a measure of the variability of our results, i.e. a measure of the spread of our distribution. The formula used is $s= [(\sum (X-M_X)^2/(N-1)]^{1/2}$, where M_X is the mean of the sample, X is the value of each measurement and N is the number of samples. Therefore, the errors shown are not the error bars on our estimation of the mean values; instead, they are the standard deviations of the distributions of experimental data points obtained for each measurements. The absorbance ratio and concentration numbers used for the calculations were the ones given by the spectrophotometer (i.e. three digits after the decimal). However, only the two first significant digits after the decimal were kept for the tables, unless the number was lower than 0.005 (in the case of the concentration in μg/μL).

2.7 Sequencing of purified plasmid DNA

Sequencing tests were mostly run on an ABI 377 sequencer (Applied BioSystems), but some of the samples were run on an ABI 310 sequencer (Applied Biosystems). Sequencing of pure plasmid was always used as a control for sequencing. The sequencing reactions and PCR were done according to the ABI Prism BigDye Terminator Cycle Sequencing Ready Reaction Kits (Original & Version 2.0) instructions and the primers used were M13 forward or M13 reverse primers (5 μM; Life Technologies). 6 μL of Big Dye Terminator Ready Reaction mix (Applied Biosystems) was added to the DNA (between 200 ng and 500 ng for a maximum of 10 μL of DNA) and the primer (0.64 μL of M13 forward or reverse primer; Life Technologies). The total volume of the reaction was 20 μL and was adjusted with deionized water. The sequencing PCR was carried out on a Progene thermal cycler (Mandel Scientific) at 96°C for 2 min, followed by 25 cycles at: 96°C for 10 s, 50°C for 5 s and 60°C for 4 min. The reactions were cleaned with a Sephadex G-50 spin column (Amersham Pharmacia Biotech) and dried down in a vacuum centrifuge.

2.8 Pure plasmid electrofiltration

The pure plasmid electrofiltrations followed the same steps (section 3.1-3.2), except that the sample added to the loading chamber was a 130 μL solution of pure plasmid solution (¼ pBluescript SK + II (0.376 μg/μL), ¼ ddH20 and ½ 0.5× TAE). The reference for the spectrophotometer readings was 0.5× TAE for both the loading and recovered samples.

2.9 Chromatographic purification

The results shown in Table 10a) were obtained following the electrofiltration protocol described in section 2.3.2, with the modification that half of the electrofiltrations were done with a 10MD polycarbonate middle membrane to obtain a more concentrated sample (vs the sample obtained while using the 14kD ×2 regenerated cellulose membranes). The samples recovered from the ten electrofiltrations (5 with a 10MD polycarbonate middle membrane, 5 with a 14kD ×2 regenerated cellulose middle membrane) were pooled together to obtain two 500 µL samples. The concentration (µg/µL) and absorbance ratio (A_{260}/A_{280}) of these two samples were evaluated with a spectrophotometer, after which they were concentrated in a vacuum centrifuge to a final volume of 20 µL. Both 20 µL samples were used as a replacement for the bacterial pellet in the chromatographic purification done with the Qiagen mini-prep column kit. As a control, mini-prep purification of a culture (standard protocol) was also done, as well as the mini-prep purification of 30 µL of pure plasmid pBluescript SK + II (0.376 µg/µL). The chromatographic purification was done according to the QIAprep Miniprep Handbook (Qiagen).

2.10 Proteinase K treatment

The electrofiltrated recovered plasmids were also treated with proteinase K to remove any proteins that could be present in the solution or complexed with the plasmids. Samples (500 µL of recovered plasmid) were generated as described for the chromatographic purification in section 2.8 (above), except that instead of using a 10MD polycarbonate, a 50kD polycarbonate was used as the middle membrane for five of the

electrofiltrations (the other five were done with 14kD ×2 regenerated cellulose). The 14kD ×2 regenerated cellulose and 50kD polycarbonate samples (500 µL each) were incubated with 1.25 µL of proteinase K (20 mg/mL) for 30 minutes at 37°C to degrade any proteins that could have been present in the solution. After the proteinase K treatment, the enzyme was inactivated by adding 500 µL of phenol, followed by a 10 s vortex and one minute centrifugation. The supernatant was transferred in a new tube with the addition of 500 µL of chlorophorm/isoamyl alcohol 24:1. The solution was vortexed for 10 s and centrifuged for 1 min. The supernatant was again transferred to another tube and 50 µL of sodium acetate (3M) and 1000 µL of ethanol at –20°C were added. The tube was centrifuged for 15 minutes, the ethanol was removed, an additional quick spin was done, the remaining ethanol was discarded, and the plasmid DNA was resuspended in 0.5× TAE

CHAPTER 3. Results

3.1 Boiling lysis buffer: STET at different concentrations

The preliminary experiments were done using a protocol with STET 1, i.e. the STET solution described in the original protocol (Holmes and Quigley 1981). When we decided to skip the preliminary centrifugation to obtain the bacterial pellet and to lyse directly the cells in the culture, we decided to use increasingly concentrated STET solutions (STET 2, STET 4 and STET 8) in order to have a lysis buffer as concentrated as possible (to get as close as possible to the original ratio of cells to STET solution). Increasing the concentration of the buffer also allowed us to have a larger culture/lysis buffer ratio, i.e. it allowed us to use a larger volume of culture relative to the volume of STET and thus maximizing the number of bacterial cells used in each experiment. Similar results were obtained with all STET buffers (not shown). Therefore, we used the STET 8 buffer in all experiments below because it allowed us to use a larger volume of cell culture.

3.2 Conditions of electrofiltration

3.2.1 Time of electrofiltration

The time duration of the electrofiltration affects the quality (yield and purity) of the recovered plasmid DNA. A number of tests were therefore performed to assess the electrofiltration time that would generate the largest amount of plasmid DNA and the highest purity ratio. Electrofiltration times of 10, 20, 25, 30 and 40 minutes were used

with the same bacterial lysate. A sample of the loading and recovery chambers after each electrofiltration were collected and run on an agarose gel to assess the amount of plasmid DNA in each of these chambers. Figure 3 shows that a maximum amount of plasmid DNA has already cross the middle membrane after 10 minutes, and that no more gain is achieved if the electrofiltration is carried out for a longer period of time. Similar experiments were done with a different middle membrane, 14kD ×2 regenerated cellulose, and the results of the spectrophotometer readings of the recovered samples are shown in Table 1. After 10 minutes of electrofiltration, both the recovery and the purity of the recovered plasmid DNA are maximized. When the electrofiltration is carried out for 20, 30 and 40 minutes instead, the amount of plasmid DNA does not increase. In fact, the amount of plasmid DNA has a tendency to slightly decrease, probably due to some loss of plasmid DNA through the end membranes.

The electrofiltration also includes a terminal 30 seconds backward pulse at 250V after the 10 minutes of electrofiltration are finished. This backward pulse serves to release any plasmid DNA that might have accumulated on the end membrane (6-8kD regenerated cellulose) of the recovery chamber. Different backward pulse durations were tried to determine what minimum time was necessary, and whether this backward pulse was useful to increase the amount of DNA recovered. Figure 4 shows that the addition of a backward pulse at the end of the electrofiltration increases the amount of plasmid DNA collected in the recovery chamber, and that a 30 second backward pulse is sufficient to release any plasmid DNA stuck on the end-membrane surface of the recovery chamber during electrofiltration.

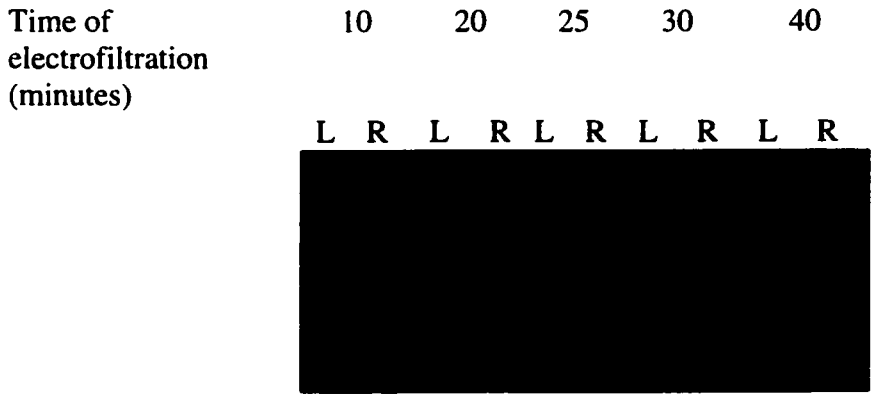


Figure 3. Amount of plasmid DNA recovered vs time duration of the electrofiltration with a 300kD PC middle membrane. The gel is loaded with an aliquot of the loading and recovery wells obtained after 10, 20, 25, 30 and 40 minutes of electrofiltration (L: loading well, R: recovery well).

Table 1. Electrofiltrations with a 14kD ×2 regenerated cellulose middle membrane for 10, 20, 30 and 40 minutes. Recovered samples were measured for concentration (μg/μL) and absorbance ratio. Values in parentheses are the standard errors and N is the number of samples electrofiltrated.

Electrofiltration time (minutes)	N	Ratio A_{260}/A_{280}	Concentration (μg/μL)
10	4	1.80 (0.05)	0.04 (0.002)
20	4	1.74 (0.06)	0.03 (0.003)
30	3	1.91 (0.12)	0.03 (0.003)
40	2	2.44 (0.37)	0.03 (0.01)

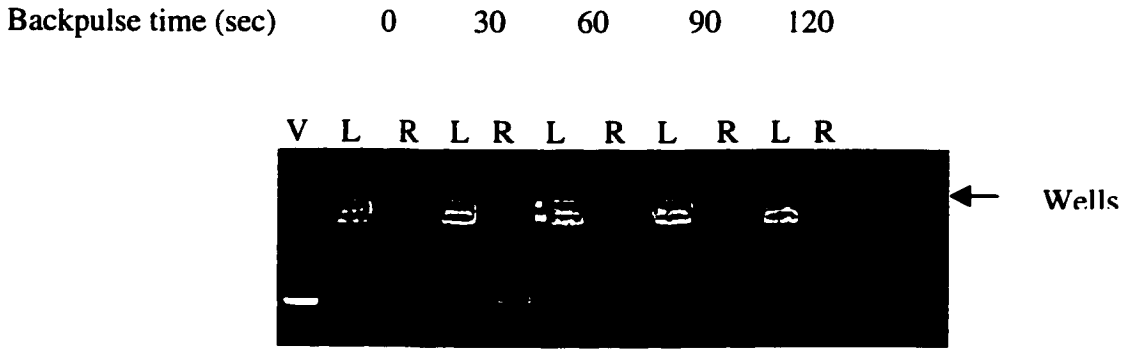


Figure 4. Amount of plasmid DNA present in the recovery chamber after 10 minutes of electrofiltration followed by a 0, 30, 60, 90 and 120 second backward pulse at 250V. Lane V shows the pure vector standard while lanes L are DNA present in the loading chamber after electrofiltration + backward pulse and lanes R are plasmid DNA present in the recovery chamber after electrofiltration + backward pulse.

3.2.2 Buffer: TAE vs TBE

Two running buffers are commonly used in agarose gel electrophoresis, TAE (Tris acetate EDTA) and TBE (Tris boric acid EDTA). The same two buffers were used as electrophoresis buffers and loading and recovery buffers (in the loading and recovery chambers) to electrofiltrate the plasmids in the bacterial lysate (after boiling lysis of the culture). Both buffers were used at 0.5× concentration (0.5 × TAE is 0.02 M Tris-acetate, 0.0005 M EDTA and 0.5 × TBE is 0.0445 M Tris-borate, 0.0445 M boric acid, 0.001 M EDTA) and the best one was then used at different concentrations to determine the parameters maximizing the yield and purity of plasmid DNA. Table 2 shows that using TAE buffer results in higher quality DNA (ratio of A_{260}/A_{280} 0.61 vs 0.45 for TBE), but lower concentration (0.13 μg/μL vs 0.29 μg/μL for TBE). Since, for PCR sequencing, purity is more important than yield, we therefore choose 0.5 × TAE as the electrophoresis buffer in all subsequent experiments.

3.2.3 Concentration of buffer (TAE)

The electrofiltration buffer can be used at different concentrations, and five different TAE concentrations were tested. Samples of bacterial lysate were electrofiltrated in electrophoresis buffers of 0.1 ×, 0.2 ×, 0.3 ×, 0.5 × and 1.0 × TAE. The results shown in Table 3 indicate that the concentration of the buffer has some effect on the amount of DNA recovered, and that the 0.5 × TAE buffer is the one that yields the most plasmid DNA with the best purity ratio. Although the results at 0.5 × TAE were not significantly different from at least one of the other conditions (0.2 × TAE) and that the

Table 2. Average concentration (μg/μL) and absorbance ratio of recovered samples after electrofiltrations with two different buffers, 0.5 × TAE and 0.5 × TBE and a permissive middle membrane (300kD polycarbonate). Each buffer was used as the running buffer and as the loading and recovery buffer in the chambers. Values in parentheses are the standard error and N is the number of samples electrofiltrated.

Buffer	N	Ratio A_{260}/A_{280}nm	Concentration (μg/μL)
0.5 × TAE	10	0.61 (0.10)	0.13 (0.04)
0.5 × TBE	11	0.45 (0.07)	0.29 (0.14)

Table 3. Average concentration (μg/μL) and absorbance ratio of the recovered samples after electrofiltrations were performed with a 25kD CA middle membrane at 0.1 ×, 0.2 ×, 0.3 ×, 0.5 × and 1.0 × TAE. Each buffer was used as the electrophoresis buffer as well as the buffer in the recovery well. Values in parentheses are the standard error and N is the number of samples electrofiltrated.

TAE Concentration	N	Ratio A_{260nm}/A_{280nm}	Concentration (μg/μL)
0.1 ×	2	1.39 (0.27)	0.03 (0.001)
0.2 ×	2	1.61 (0.02)	0.04 (0.01)
0.3 ×	2	1.36 (0.33)	0.05 (0.00)
0.5 ×	4	1.65 (0.05)	0.05 (0.002)
1.0 ×	2	1.43 (0.09)	0.04 (0.004)

number of samples per condition was only ranging from 2 to 4, we choose 0.5 × TAE as the electrophoresis buffer in all subsequent experiments.

3.2.4 Membrane compositions and pore sizes

The most crucial element in the electrofiltration seems to be the membranes, especially the middle membrane which has the role of separating plasmid DNA from contaminants by letting the plasmids go through (and into the recovery chamber) while retaining the contaminants in the loading chamber. Whereas the pores of the end membranes have to be small enough to retain the plasmids inside the chambers during electrofiltration, the pores of the middle membrane have to be big enough to allow plasmid DNA to move into the recovery chamber but at the same time, small enough to retain as much contaminants as possible in the loading chamber. The first membranes that we used were polycarbonate membranes of various sizes. All of these membranes (with pore sizes ranging from 30kD to 10MD) can be classified as having large pores and to be permissive membranes. These membranes were used to determine most of the optimal parameters during electrofiltration; their permissive nature made it easier to differentiate between conditions. However, these membranes did not allow us to obtain very pure DNA, as judged by the spectrophotometer readings, so other types of membranes were tested, such as cellulose acetate membranes. These membranes are also available in a variety of pore sizes (25kD, 100kD, 300kD), but all of them were more restrictive, i.e., a smaller amount of both contaminants and plasmid DNA could pass through them. Table 4 shows a comparison between same pore size (300kD) membranes of polycarbonate and cellulose acetate. As mentioned previously, the polycarbonate membranes allow more

Table 4. Average concentration (μg/μL) and absorbance ratio of the recovered sample after electrofiltrations performed with 2 different 300kD middle membranes, polycarbonate and cellulose acetate. Values in parentheses are the standard errors and N is the number of samples electrofiltrated.

Membranes	N	Ratio A_{260}/A_{280}	Concentration (μg/μL)
Polycarbonate	7	0.59 (0.07)	0.12 (0.04)
Cellulose acetate	7	1.09 (0.16)	0.09 (0.04)

DNA to be recovered (0.12 μg/μL) than with the cellulose acetate membranes (0.09 μg/μL), but it is of lower quality (A_{260}/A_{280} ratio of 0.59 vs 1.09 for cellulose acetate).

Finally, since even the samples obtained using the cellulose acetate membranes could not be sequenced (and the absorbance ratio A_{260}/A_{280} was still not at 1.800), a 14kD regenerated cellulose membrane was used as the middle membrane. In fact, two layers of this membrane were used in order to minimize the amount of contaminants migrating to the recovery chamber. Table 5 shows a comparison study of eleven membranes, including six polycarbonate membranes ranging from 30kD to 10MD, 3 cellulose acetate membranes from 25kD to 300kD and two regenerated cellulose membranes, a 14kD and a double layer of 14kD membranes (14kD ×2). The results in Table 5 clearly demonstrate that the nature of the membranes - polycarbonate, cellulose acetate or regenerated cellulose - has a much larger effect on the amount and purity of the recovered plasmid DNA than the size of the membranes. Furthermore, membranes of similar compositions (cellulose acetate and regenerated cellulose) gave similar results. Since our goal is to do sequencing, the best choice of membranes is the double layered 14kD regenerated cellulose because it gave the cleanest plasmid DNA (as judged from the A_{260}/A_{280} ratio). Though it does not provide plasmid DNA that is significantly different from the 14kD single layer, the 14kD ×2 was the middle membrane chosen for electrofiltration in hope that it would block more contaminants that could be interfering with the sequencing and that may not be picked up by the spectrophotometer.

In order to take advantage of the best qualities of both small and big pore membranes, combinations of middle membranes were also tested for electrofiltration (1

Table 5. Comparison of different middle membranes for electrofiltration according to the average concentration (μg/μL) and absorbance ratio of the recovered samples. Values in parentheses are the standard errors and N is the number of samples electrofiltrated. RC: regenerated cellulose, CA: cellulose acetate, PC: polycarbonate

Membranes	N	Ratio A_{260}/A_{280}	Concentration (μg/μL)
14kD ×2 RC	5	1.83 (0.02)	0.04 (0.01)
14kD RC	12	1.80 (0.30)	0.05 (0.01)
25kD CA	12	1.71 (0.36)	0.05 (0.02)
30kD PC	9	0.70 (0.15)	0.09 (0.04)
50kD PC	11	0.93 (0.47)	0.07 (0.03)
100kD CA	12	1.63 (0.27)	0.07 (0.02)
300kD CA	12	1.63 (0.17)	0.07 (0.02)
300kD PC	12	0.81 (0.35)	0.17 (0.10)
500kD PC	11	0.72 (0.15)	0.19 (0.04)
2MD PC	11	0.74 (0.20)	0.21 (0.09)
10MD PC	11	0.73 (0.18)	0.22 (0.07)

big pore membrane combined to 1 small pore membrane). The small pore membranes were the 14kD and the 50kD regenerated cellulose membranes while the big pore membranes were polycarbonates membranes with pores ranging from 300kD to 10MD. Table 6 shows that these electrofiltrations were not successful in increasing the amount of plasmid DNA recovered without decreasing the A_{260}/A_{280} ratio.

3.3 PCR dilution tests

As a first mean to assess the purity of the recovered plasmid DNA, we first did PCR dilution tests on the recovered samples to see if there were contaminants that could prevent PCR. Many samples were used in PCR reactions at different concentrations (ex: 1×, 2×, 10×, 20×, 40×, 80×) to see if the PCR signal decreased with less DNA in the reactions, or if the PCR signal increased rather than decreased following the dilution of the samples. If the PCR signal decreases with increasing dilutions, it means that the plasmid DNA did not contain contaminants that prevented PCR; on the other hand, if the PCR signal shows an increase in signal with increasingly diluted samples, it means that there is contamination present, and that the dilution of that contaminant (along with the dilution of DNA) allows for better PCR, and therefore increased PCR signal. Figure 5 illustrates that the intensity of the PCR products of increasingly diluted electrofiltrated plasmid DNA samples (1×, 10×, 20×, 40×, 80×) decreases with every dilution, indicating that there are no PCR-inhibiting contaminants present.

3.4 Sequencing results of purified samples

The goal of our project is to produce sequencing grade plasmid DNA. Since PCR is much more tolerant to contaminations and impurities than PCR sequencing, the electrophoresed samples were thus subjected to PCR sequencing. Unfortunately, the numerous samples subjected to sequencing all failed to produce any sequence, even though the amount of DNA was sufficient and the purity was optimal (1.800) according to the spectrophotometer values. These findings led us to investigate why the samples could not be sequenced, and if any contaminant could be coming from the preparation of the boiling lysate or the electrofiltration itself. Pure plasmid was used to carry out these controls, with the addition of various chemicals or manipulations (or a mix of both) that came across during the isolation and purification process of our electrofiltrated plasmid. Table 7 lists the controls that were employed to identify any condition that could be inhibiting the sequencing of the recovered plasmid. As seen in Table 7, the presence of TAE and STET in the sequencing reactions does not affect it, nor does it affect the reaction if these two products are boiled or electrofiltrated prior to the sequencing. Moreover, passage through the middle membrane (50kD PC) itself does not affect the plasmid in any way that could prevent sequencing, since the electrofiltrated pure plasmid was sequencable.

Table 6. Comparison of different combinations of middle membrane of small and big pores with the average concentration (μg/μL) and absorbance ratio of the recovered samples. Values in parentheses are the standard errors and N is the number of samples electrofiltrated. RC: regenerated cellulose, CA: cellulose acetate, PC: polycarbonate

Membranes		N	Ratio A_{260}/A_{280}nm	Concentration (μg/μL)
Small pores	Big pores			
14kD RC	300kD PC	2	1.99 (0.27)	0.02 (0.00)
14kD RC	500kD PC	4	1.67 (0.10)	0.04 (0.004)
50kD RC	300kD PC	2	1.91 (0.03)	0.02 (0.004)
50kD RC	500kD PC	4	1.62 (0.09)	0.04 (0.003)
14kD RC	2MD PC	2	1.65 (0.001)	0.05 (0.004)
14kD RC	10MD PC	2	1.66 (0.05)	0.05 (0.001)
50kD RC	2MD PC	2	1.62 (0.01)	0.05 (0.002)
50kD RC	10MD PC	2	1.58 (0.06)	0.05 (0.003)

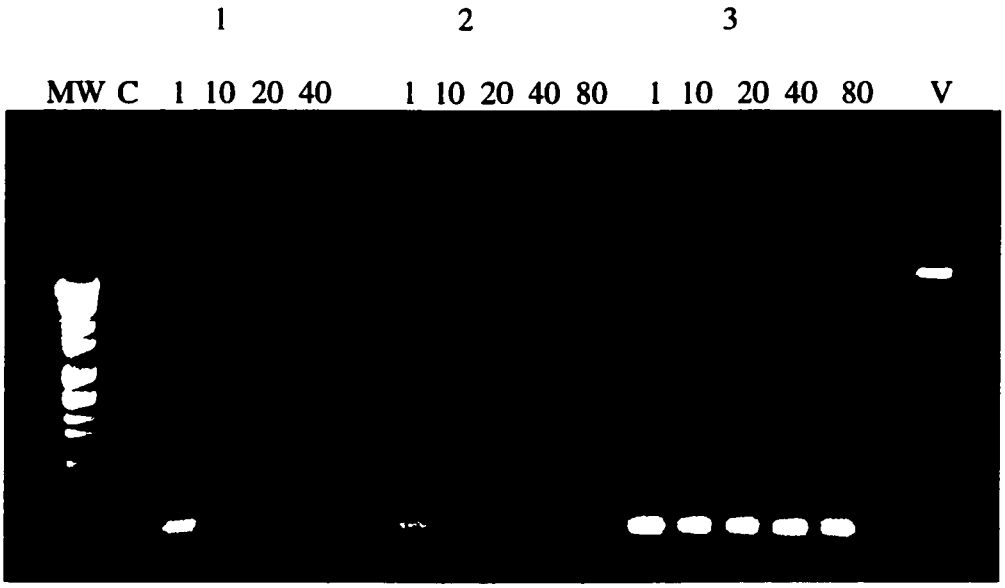


Figure 5. PCR dilution tests on 3 different samples electrofiltrated with 300kD PC (1), 500kD PC (2) or 2MD PC (3) middle membranes. Each recovered sample was diluted 1×, 10×, 20×, 40× and 80× and PCR was performed on these diluted samples.

C: negative control, V: pure vector (no PCR), MW: molecular weight marker

3.5 Pure plasmid electrofiltration

In order to obtain the electrofiltrated pure plasmid mentioned in the section above, pure plasmid was first electrofiltrated with two 14kD regenerated cellulose middle membranes, since that was our chosen condition of electrofiltration. Surprisingly, essentially no plasmid could be recovered when we performed such an electrofiltration. Absorbance measurements of the plasmid DNA present in the loading and recovery chambers (following electrofiltration), showed that most of the pure plasmid DNA had not migrated to the recovery chamber and that it had stayed in the loading chamber (see Table 8). Due to these unexpected results, a series of electrofiltrations of pure plasmid was done with membranes of various nature and size. Table 8 clearly shows that pure plasmid does not cross the middle membrane if this membrane had small pores or was not permissive (see classification in the section 3.2.4). The membrane that allowed the most plasmid to move to the recovery chamber was the 50kD polycarbonate membrane. The quasi absence of the plasmid DNA in the recovery chamber in negative cases was not due to its loss: between 75% and 88% of the plasmid DNA could be recovered from the loading and the recovery chambers. Moreover, the results of the electrofiltration experiments where we varied the time durations are consistent with our previous results: the amount of DNA that crosses the middle membrane increases if the electrofiltration time is increased (Table 9). Table 9 shows an increase of plasmid DNA in the recovery chamber between 10 and 20 minutes of electrofiltration, and between 20 and 30 minutes. Contrary to the experiments done with boiling lysate samples, where the amount of plasmid DNA recovered plateaued after 10 minutes of electrofiltration (probably due to the clogging of the middle membrane by the contaminants present in the lysate), the pure

plasmid is not submitted to the same plateau effect because there are no contaminants present in the loaded sample that could potentially clog the middle membrane.

3.6 Physical properties of electrofiltrated plasmid

To ensure that the plasmids were still supercoiled after the electrofiltration (in order to rule out this as the reason why PCR sequencing was not possible), boiling lysates were electrofiltrated under 2 conditions, one with a permissive middle membrane (300kD PC) and one with a restrictive middle membrane (14kD ×2 RC). Figure 6 a) shows a supercoiled plasmid after electrofiltration with the 300kD PC membrane (lane 3). The supercoiled nature of the plasmid can be determined because it runs at the same position as the pure vector control (which is supercoiled) in lane 1 and as the plasmid from the boiling lysate in lane 2. The two upper bands visible in lane 2 and 3 are the two other forms of plasmid present in the recovered samples, the highest band being the open circular form, and the middle band being the linearized form of the plasmid. In figure 6 b), a similar result is observed, this time with a 14kD ×2 RC membrane. The electrofiltrated plasmid in his supercoiled form is in lane 3, whereas the pure vector control is in lane 1 and the culture lysate is in lane 2. Since the plasmids remain supercoiled after electrofiltration, this cannot explain the failed sequencing attempts.

Table 7. List of various controls for sequencing reactions with pure plasmid (PP).

Samples	Sequencing results	Length of sequence read (bases)
Pure plasmid (PP)	+	690
PP + TAE (low concentration 0.05×)	+	660
PP + TAE (high concentration 0.2×)	+	650
PP + STET8 + 5min. at 65°C	+	540
PP + STET8 + 5min. at 100°C	+	480
PP + electrofiltrated TAE	+	480
PP + electrofiltrated STET4 and 0.5 TAE	+	640
PP + electrofiltrated STE4 and 0.5 TAE	+	640
PP electrofiltrated (50kD PC)	+	450

Table 8. Average concentration (μg/μL) and absorbance ratio of pure plasmid samples collected from the loading and recovery wells after electrofiltration with different membranes. Values in parentheses are the standard error and N is the number of samples electrofiltrated.

Membranes	N	Loading chamber			Recovery chamber		
		Ratio A_{260}/A_{280}	Concentration (μg/μL)	% of total sample	Ratio A_{260}/A_{280}	Concen-tration (μg/μL)	% of total sample
14kD x2 (RC)	2	1.80 (0.01)	0.07 (0.001)	75%	0.35 (0.30)	-	-
30kD (PC)	4	1.82 (0.02)	0.03 (0.01)	34%	1.80 (0.01)	0.04 (0.01)	47%
50kD (PC)	4	1.79 (0.02)	0.016 (0.01)	16%	1.98 (0.02)	0.06 (0.01)	66%
50kD (CA)	2	1.81 (0.003)	0.07 (0.004)	77%	1.92 (0.28)	0.01 (0.01)	11%
100kD (CA)	1	1.82	0.07	79%	1.98	-	-
300kD (CA)	2	1.81 (0.001)	0.07 (0.001)	69%	1.80 (0.001)	0.01 (0.002)	7%
300KD (PC)	4	1.80 (0.01)	0.04 (0.01)	44%	1.79 (0.01)	0.03 (0.01)	35%
500kD (PC)	3	1.81 (0.004)	0.07 (0.003)	74%	1.80 (0.05)	0.01 (0.002)	10%
10MD (PC)	4	1.80 (0.01)	0.07 (0.01)	70%	1.79 (0.05)	0.01 (0.003)	10%

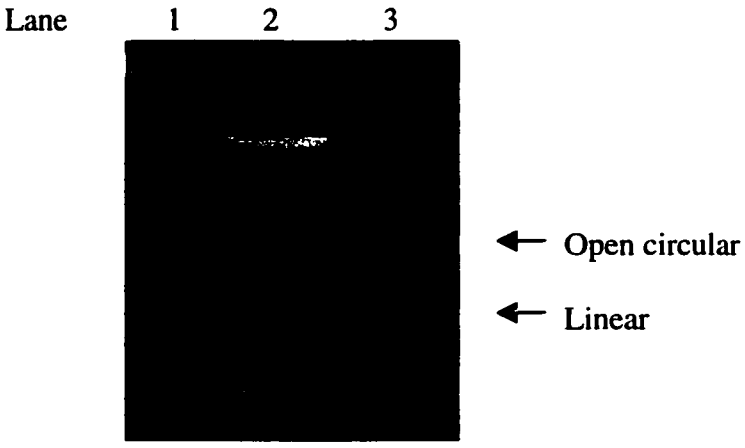
Notes: -: not enough DNA to be measured

Table 9. Average concentration (μg/μL) and absorbance ratio of pure plasmid from the loading and recovery wells after electrofiltration for 10, 20 and 30 minutes. The middle membrane is a 50kD PC. Values in parentheses are the standard errors and N is the number of samples electrofiltrated.

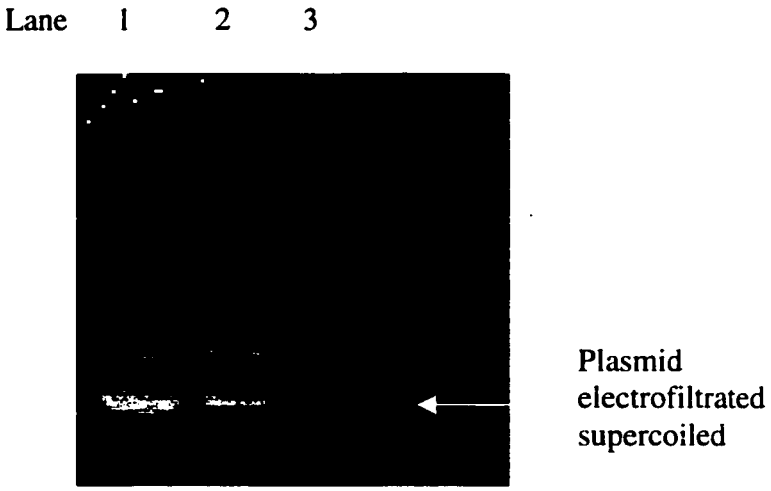
Time of electro-filtration	N	Loading chamber			Recovery chamber		
		Ratio A_{260}/A_{280}	Average conc. (μg/μL)	% of sample	Ratio A_{260}/A_{280}	Average conc. (μg/μL)	% of sample
10 min.	4	1.80 (0.01)	0.04 (0.02)	44%	1.80 (0.01)	0.03 (0.01)	36%
20 min.	4	1.83 (0.02)	0.02 (0.01)	20%	1.80 (0.004)	0.06 (0.01)	61%
30 min.	4	1.98 (0.05)	0.01 (0.001)	12%	1.85 (0.09)	0.080 (0.01)	82%

3.7 Chromatography of electrophoresed samples

In light of the previous results on sequencing and electrofiltration of pure plasmid, some further investigations were made because our data showed that only the pure plasmid could not cross the membrane, while only the electrofiltrated plasmid from a culture could not be sequenced. In the first set of experiments, we purified the electrofiltrated samples with a standard chromatographic column, in this case a QIAGEN column. Samples were electrofiltrated both with a restrictive and a permissive middle membrane to obtain a certain amount of plasmid DNA, one at low concentration (Table 10a, 14kD ×2 RC) and one at high concentration (Table 10a 50kD PC). These samples were concentrated from a volume of 500 µL to a volume of 20 µL to respect as much as possible the protocol guidelines. No plasmid could be recovered from these samples after chromatographic purification, while a culture and the pure plasmid did produce pure DNA (see table 10b), again underlining the differences in properties between a pure plasmid and our recovered plasmid.



a)



b)

Figure 6. Form of the plasmids after electrofiltration. Electrofiltrated plasmids are still supercoiled after passing through a middle membrane of a) 300kD PC (lane 3) and b) 14kD ×2 RC (lane 3). In both Figure 6 a) and b), lanes 1 and 2 are the pure plasmid and the boiling lysate, respectively.

Table 10 a) Concentration (μg/μL) of samples recovered after N=10 electrofiltrations with 14kDx 2 regenerated cellulose or a 10MD polycarbonate middle membrane. These samples have been used for the chromatographic purification (see table10b).

Membranes (middle)	# of samples electrofiltrated	Total μL	Ratio A_{260}/A_{280}	Concentration (μg/μL)	Amount of DNA (μg)
14kDx2 RC	5	500	1.69	0.02	9.5
10MD PC	5	500	1.38	0.10	51.5

Table 10 b) Concentration (μg/μL) and absorbance ratio of samples that were chromatographically purified with a Qiagen column.

Samples	μL (concentration μg/μL)	Ratio A_{260}/A_{280}	Concentration (μg/μL)	μg of DNA	% of DNA recovered
Culture	575 μL	1.80	0.05	4.9 μg	NA
Pure plasmid	30 μL (0.376)	1.79	0.06	6.3 μg	55.9%
14kDx2 RC	500 μL reduced to 20 μL (0.019)	1.73	0.002	0.2 μg	2%
50kD PC	500 μL reduced to 20 μL (0.103)	1.82	0.002	0.2 μg	0.4%

3.8 Proteinase K purification of electrofiltrated samples

In the second set of experiments, we further treated the recovered plasmid DNA using proteinase K. The goal was to remove any protein that could be attached to the electrofiltrated plasmid DNA and that could change its properties in a way that would allow it to cross the middle membrane while also preventing PCR sequencing. Table 11a shows two samples electrofiltrated with a restrictive (14kD ×2 RC) and a permissive membrane (50kD PC). These samples were both treated with proteinase K according to the protocol described in the section 2.9 and the recovered DNA was evaluated with a spectrophotometer, giving the results shown in Table 11b. Although this extra purification step should have increased the purity (absorbance ratio of A_{260}/A_{280}), it was not the case. In fact, the purity seems to have decreased after the proteinase K treatment. The residual presence of phenol in the samples could explain this decrease in purity. The 14kD ×2 RC sample did not provide enough DNA to perform a sequencing reaction, but the 50kD PC sample did and it gave positive sequencing results, indicating that a protein could have been removed from the sample and that that protein was preventing PCR sequencing of the plasmid DNA. Therefore, we can say that there is a protein attached to the plasmid, and that this protein is necessary for the electrofiltration, but that it also prevents sequencing. Our method would then call for proteinase K, but this additional step makes our method uncompetitive.

Table 11. Concentration (μg/μL) and absorbance ratio of recovered samples electrofiltrated with a 14kD ×2 regenerated cellulose or a 50kD polycarbonate membrane before and after being treated with proteinase K.

a) After electrofiltration, before proteinase K treatment.

Membranes (middle)	# of samples electrofiltrated	Ratio A_{260}/A_{280}	Concentration (μg/μL)	μL of DNA
14kD x2 RC	5	1.67	0.04	500
50kD PC	5	1.38	0.12	500

b) After proteinase K treatment

Membranes	Ratio A_{260}/A_{280}	Concentration (μg/μL)	Sequencing
14kD x2 RC	1.04	0.01	NA
50kD PC	1.23	0.05	Positive sequencing

CHAPTER 4. Discussion

Our goal was to create a plasmid isolation and purification system that would be as streamlined and automation-friendly as possible. This required to remove all the centrifugation steps because centrifugation is the hardest step to automate. Most protocols on the market actually require one or more centrifugations in their protocols. In some of the commercially available systems, the centrifugation steps have been replaced by vacuum filtration steps. However, the robots that allow the automatisation of these protocols are still very expensive. Our goal was therefore to develop a centrifugation-free system that would be easy to automate. We hypothesized that a purification system based on electrophoresis could be developed to meet this goal.

4.1 Electrofiltration: Effect of different parameters

The results indicate that the TAE buffer provides the highest purity of recovered plasmid when compared to the recovered plasmid produced during electrofiltration with a TBE buffer. Although the TBE recovered plasmid is more concentrated, using the TAE buffer produces a recovered plasmid of higher purity and sufficient concentration for sequencing. It is possible that the salt composition of the buffer affects the surface charge of the middle membrane. It has been shown that TBE ions have better buffering capacity and that they move more slowly than TAE ions during electrophoresis (Desruisseaux *et al.* 1998). This could imply that ions accumulate more slowly onto the middle membrane surface during electrofiltration. This lower ion accumulation on the membrane could explain why more plasmid is able to pass through the membrane with a TBE buffer

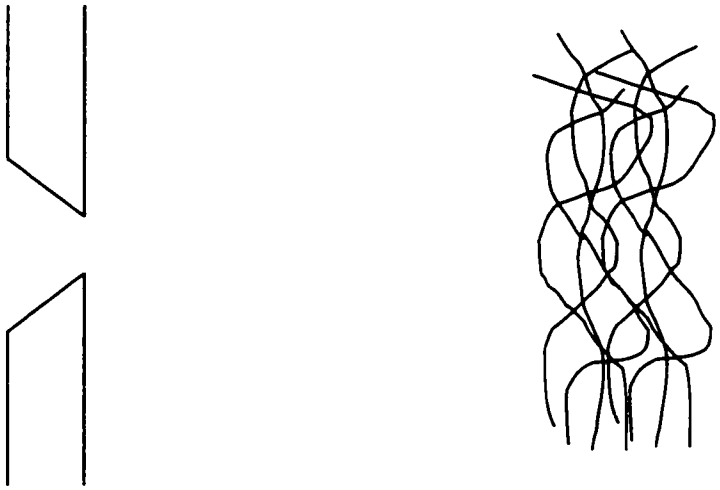
(although along with more contaminants). Indeed, less ion accumulation would mean less repulsive forces between the middle membrane and the plasmids. As for TAE, due to its faster ionic migration, more ions could accumulate on the middle membrane surface, therefore creating more repulsive forces between the plasmids (and the contaminants) and the middle membrane.

As seen in the results, the parameter that has the most effect on the recovered plasmid regarding both concentration and purity is the middle membrane, because it is the major element that separates plasmid DNA from all the other contaminants. Moreover, it is in fact the nature of the membrane and not the pore size of this one that really determines how much plasmid DNA will go through and how pure it will be (or how much contaminants will be prevented from moving to the recovery chamber along with the plasmids). As described in the results, the membranes were classified in to two groups, the restrictive membranes (allow less DNA to go through, but also less contaminants) and the permissive membranes (allow more DNA to go through, but also more contaminants). These membranes can be put in the same two groups according to the nature of the materials there made of, i.e. the restrictive membranes were all made of cellulose (regenerated cellulose or cellulose acetate) and the permissive ones were all made of polycarbonate.

The pore size of the membranes, which should have determined how much DNA and contaminants were going through, did have a small effect on these values. When looked at on a large scale of pore sizes, like for the polycarbonate membranes, there is some effect due to the pore size of the membrane, but it is definitively less important than the membrane composition effect. There are more differences in the concentration and

the ratio of the plasmid recovered from two 300kD membranes, one cellulose acetate and one polycarbonate, than between a 300kD polycarbonate membrane and a 10MD membrane of the same composition. These differences could be due to how the membranes are made, or more precisely how the pores are made: the polycarbonate membranes are flat surfaces that are punched through using particle bombardment, making precise holes with a very flat surface (Figure 7a). As for the cellulose acetate and regenerated cellulose, the surface is far from being flat since the membrane is made of a mesh of fibers that are all entangled and cross-linked together (Figure 7b). The pore size is then estimated as the average pore size created by the mesh, depending on how tight the mesh is. This composition and physical arrangement is probably much more effective in trapping contaminants of various sizes, but it also traps more DNA, letting less DNA move to the recovery chamber. In parallel, we can observe that the electrofiltration results for the same conditions (with a 300kD polycarbonate membrane) in Table 4 and Table 5 are very different. These differences are due to the large day-to-day variation between the cultures (OD of culture) used for each set of electrofiltrations.

The fact that significant differences are hard to see among same composition membranes of different pore sizes is probably due to the rapid clogging of the membranes. The concentration of contaminants in the boiling lysate is important enough that clumps of cellular debris and genomic DNA/proteins/SDS aggregates are visible in the boiling lysate (the solution is cloudy), and these contaminants are also pipetted in the loading well. It has been observed that after 10 minutes of electrofiltration, there is already a film of solid material that has accumulated on the loading side of the middle



a) b)

Figure 7. a) Schematic representation of a polycarbonate pore and the surface of the membrane. b) Schematic representation of a cellulose pore and the surface of the membrane

membrane, probably blocking most of the pores and preventing more DNA from moving into the recovery chamber. The thirty second backward pulse in not strong enough to remove that film from the membrane, nor is a longer backward pulse. This likely explains why no significant gain of plasmid DNA is observed if electrofiltration is carried on for longer durations, and why this phenomenon is not observed when electrofiltration (with a permissive membrane) is done with a pure plasmid solution. Although pure plasmid does not cross restrictive and most permissive membranes, it can cross a 50kD polycarbonate membrane, as well as the 30kD polycarbonate and 300kD polycarbonate membranes, though in smaller quantities. The amount of pure plasmid DNA that can pass the middle membranes mentioned above actually reaches a maximum value with the 50kD polycarbonate membrane. Surprisingly, pure plasmid DNA could not cross polycarbonate membranes with larger pore sizes, such as the 500kD and 10MD polycarbonate membranes. The electrophoresis with the 50kD polycarbonate membrane allowed us to perform electrofiltration experiments with pure plasmid for various time durations. After ten minutes of electrofiltration, more than a third of the pure plasmid had crossed the middle membrane. Although this increase was not linear when the electrofiltration was carried on for 20 and 30 minutes, the amount of DNA that crossed the middle membrane with the extra time of electrofiltration was significant. Because the pure plasmid DNA solution introduced in the loading chamber does not contain any contaminants that could block the pores of the middle membrane or that could aggregate on its surface, more plasmid DNA is not prevented from migrating to the recovery chamber when the electrofiltration is carried on for more than ten minutes.

Finally, the electrophoresis buffer concentration does not have any observable significant effect on the recovered plasmid. This can be due to two factors: 1) The buffer may not in fact have any effect on the recovered plasmids, or 2) if it does have an effect, this effect may not be observable. Since only a low proportion of the actual amount of plasmid DNA present in the loading chamber crosses the middle membrane to the recovery chamber because of the blocked middle membrane, it could mask the influence of electrophoresis buffer on plasmid recovery. If only a small proportion of plasmid DNA present in the loading chamber can actually go through the membrane, then any effect that the buffer concentration could have will be minimized. That reduction could be enough to prevent the observation of any differences.

4.2 PCR dilution test and sequencing

4.2.1 PCR dilution test: not informative

As a first mean of assessing the purity of our recovered samples, PCR was performed on all of the recovered samples. Since PCR worked all the time on all our samples (no matter how they were produced), PCR dilution tests were then done to see if any contaminant was present in the recovered samples. As showed in section 3.3, the PCR dilution tests did not reveal the presence of contaminants preventing PCR, and we therefore wrongly assumed that the same samples could be sequenced. When we examined sequencing and realized that no positive results could be generated, we started using the spectrophotometer to assess concentration and purity of our recovered samples. These values were much more informative in determining the quality of the recovered

samples, and no more purity PCR was done because this method is not sensitive enough in this situation.

4.2.2 Sequencing does not work with electrofiltration only

Since our goal was to produce plasmid DNA that could not only be amplified, but also sequenced, PCR was replaced by the spectrophotometer to assess purity for sequencing. Although the electrofiltration parameters have been modified (small pore membranes instead of big pore membranes, STET 8 instead of STET 1) to obtained recovered plasmid DNA that should theoretically be sequencable (enough DNA to meet the Big Dye kit standard, absorbance ratio A_{260}/A_{280} close to 1.800), no samples purified by electrofiltration alone could generate reliable sequences. In most laboratories, and for most methods, the purity of DNA is assessed by measuring the ratio of absorbance at 260nm and 280nm. Although this method is usually accurate, it does not seem to be in our particular case, because even samples that had optimal absorbance ratios at 260nm/280nm were not sequencable. Our results thus show that this method is not perfect, and that there are some contaminants or molecules that cannot be detected by the spectrophotometer.

4.3 Purity of DNA purified by electrofiltration

4.3.1 STET, TAE (concentration) and boiling

Since the recovered plasmid DNA was not in a regular buffer solution (ex: H_2O or TE) after the electrofiltration but in a mixture of 0.5 X TAE, STET and LB, the chemicals that were involved in the electrofiltrations and that did end up in the recovery buffer were tested separately and in combinations to see if they would prevent sequencing of pure plasmid DNA. STET 8 was electrofiltrated by itself in TAE at the concentration used for a regular electrofiltration to mimic as closely as possible the final concentration of that solution in the recovery well. This buffer was added to the sequencing reaction at the same concentration at which it could have been found in a normal sequencing reaction. The same was done with TAE at both a low and high concentrations, and none of these additions prevented sequencing. The boiling (100°C for 5 minutes) did not have any effect on the quality of the sequenced either, which is not surprising since the sequencing PCR protocol itself includes a heating step of three minutes at 96°C before the cycling begin.

4.3.2 Further purification of DNA isolated by electrofiltration

As one of the controls for sequencing, pure plasmid was electrofiltrated to see if the electrofiltration itself was modifying the plasmid DNA in any way that could have been preventing sequencing. Surprisingly, no pure plasmid could be recovered after the electrofiltrations with the restrictive membranes, and the spectrophotometer results showed that most of the plasmid had stayed in the loading well. After a number of electrofiltrations were performed on pure plasmid, we concluded that the pure plasmid

could not cross the middle membrane under conditions where the plasmid from the boiling lysate could, except under very permissive conditions such as with a 50kD PC (as well as with 30kD PC or 300kD PC middle membranes, but in smaller amounts). The pure plasmid that could go through the 50kD PC middle membrane was used as a template to do sequencing and, as shown in the results (Table 7), the sequencing was successful. Given the important differences that seemed to exist between the pure plasmid that could not cross a restrictive membrane but could be sequenced, and the recovered plasmid that could cross a restrictive membrane but could not be sequenced, and given that all the added chemicals or transformations (TAE, STET, boiling and electrofiltration) did not seem to be preventing sequencing of a pure plasmid, further purifications were done on recovered plasmid DNA samples to investigate the subtle differences exist between a pure plasmid and a recovered plasmid.

4.3.2.1 Chromatographic purification

After eliminating a number of possibilities that could explain why the electrofiltrated recovered plasmids could not be sequenced (i.e. STET, boiling, TAE, electrofiltration), we hypothesized that some contaminants which could not be picked up by the spectrophotometer were present in the recovered plasmid solution. In order to see if those contaminants could be removed or not, the recovered plasmid DNA was re-purified with a standard method, i.e. a chromatographic column (Qiagen). As a control, pure plasmid was also re-purified, and most of that pure plasmid could be recovered after the chromatography purification. However, no electrofiltrated plasmid was recovered from the chromatography purification. This supported the hypothesis that there was a

fundamental difference between the pure and the electrofiltrated plasmids, and that this difference was not due to the presence of free contaminants in the solution, but probably to the presence of contaminants complexed with the electrofiltrated plasmid itself. The presence of these contaminants could change in some ways the properties (charge, conformation, compactness/folding) of the electrofiltrated plasmid in comparison to the pure plasmid. We thus hypothesized that a protein could be attached to the supercoiled electrofiltrated plasmid.

4.3.2.2 Proteinase K

After some samples were electrofiltrated with big pore and small pore membranes, these samples were treated with proteinase K to remove any proteins that could be attached to the plasmid DNA after electrofiltration. The samples were then phenol/chloroform extracted and ethanol precipitated to remove the proteinase K. This was done in order to insure that the proteinase K would not prevent sequencing. The recovered samples were evaluated for concentration and purity with the spectrophotometer, and this revealed that the sample generated with the small pore membrane did not had enough DNA left after the proteinase K treatment to perform sequencing, but the big pore membrane sample did. The sequencing of the sample from the big pore membrane was successful, leading us to think that there is indeed a protein attached to the electrofiltrated plasmid DNA. This attachment probably occurs during boiling lysis, and that this particular protein, or group of proteins, likely modify the size of the plasmid by modifying its folding or change the total charge of the plasmid which (remarkably) allows it to cross membranes that are much more restrictive. In contrast, the

pure plasmid that is not complexed with any proteins - it was purified with a CsCl/EtBr ultracentrifugation - cannot cross those restrictive membrane in its original supercoiled form. The protein thus acts as a carrier, but it also inhibits sequencing.

4.4 Conclusion

Our results show that a plasmid that is purified by electrofiltration cannot be sequenced, probably due to the presence of a protein that is attached to the plasmid. This protein-plasmid complex allows the plasmid to cross very restrictive membranes in comparison to the pure plasmid, but unfortunately, prevents sequencing. This protein may change in some ways the folding/compactness or the total charge of the plasmid, or both, and it is these changes that help the plasmid to go through membranes with pores small enough to prevent the migration of most contaminants. We would therefore need an extra proteinase K step that would render our method non-competitive. Because of the presence of a protein attached to our plasmid, we were not able to achieve our goal to develop a fast and robust automation-friendly method to purify plasmid DNA for sequencing. In order for the purification to work, it would be necessary to redesign a new lysis protocol that would lyse the cells, but that would not create any protein-plasmid complex. Unfortunately, given our results, it is possible to assume that a protein-free plasmid would not be able to move through the small pore membranes necessary to block enough contaminants. Even though we have reached an unexpected dead-end with the electrofiltration project, the results of the direct sequencing comparison study (Part II of this thesis) indicate that high throughput DNA sequencing may not require plasmid DNA purification after all.

Part II: Comparison of different direct colony sequencing methods and PCR-based sequencing methods

CHAPTER 5. Introduction

5.1 General introduction

As mentioned above, the need for a high-throughput plasmid purification protocol comes primarily from the advancement in sequencing methods. These new methods require large numbers of samples and the preparation of these samples can be an important limiting step.

5.1.1 Preparation of template for sequencing

As mentioned in section 1.1.2.3, the common procedure to obtain sequencing grade plasmid DNA templates for sequencing reactions is to clone the DNA into a vector, transform bacterial cells with that plasmid, grow the colony overnight in liquid media and purify the plasmid from that culture. This process is very reliable at producing high quality plasmid DNA in sufficient amounts for sequencing, but it is time-consuming since there are two overnight growths (one for the plate, one for the culture). Furthermore, most available commercial mini-preps kits are somewhat expensive, and can become labor-intensive depending on the number of samples that are to be processed at once.

5.2 Objective of the direct sequencing project

The objective of this second part is to compare different sequencing methods based on colony sequencing or PCR to assess the speed and reliability of these methods that do not require any plasmid purification. Those methods substantially decrease the

amount of time required to go from a clone to a sequence. The comparison will look at the reproducibility, quality of sequence, time and cost of six recently published methods (Brown Gladden *et al.* 2000, Sasho 2000, Silva *et al.* 2001, Smith *et al.* 2000, Truett *et al.* 2000, Zhang *et al.* 1999)

CHAPTER 6. Materials and Methods

All the methods that have been compared have for goal to do sequencing reactions for capillary electrophoresis without having to purify the plasmid DNA first. Some of these methods have been slightly modified, and each method will be described as it was executed for this comparison. The sequencing reactions were all cleaned with a Sephadex G50 column (Amersham Pharmacia Biotech) according to the manufacturer's instructions, and then dried up in a vacuum centrifuge. All the sequencing reactions were run on the same ABI 310 sequencer (Applied Biosystems) after the reactions had been resuspended in 25 μL of template suppression reagent (Applied Biosystems). In this study, a sequencing reaction was considered successful when the base calling software could read at least 200 bases.

6.1 Direct colony sequencing I: Chikako Sasho (2000): "Direct colony sequencing of plasmid DNA by dye terminator methods"

This protocol does not require any plasmid purification or culture growth to provide plasmid DNA for sequencing. The reaction was done directly on the colony, given that it had reached a certain size (to provide enough DNA). A 2-2.5mm colony was picked from the agar plate with a plastic tip that had been flamed to seal its narrow end. The colony was resuspended in 6 μL of BigDye Terminator Ready Reaction mix (ABI), to which were added 0.64 μL of M13 forward or reverse primer (5 μM), and 7 μL of ddH_2O. The reactions were carried out in a Progene (Mandel Scientific) DNA thermal

cycler at 94°C for 3 min (the step that probably frees the DNA from the cells) and then for 25-40 cycles of 96°C for 30 s; 50°C for 15 s; and 60°C for 2 min.

6.2 Direct colony sequencing II: Brown Gladden *et al.* (2000): "Rapid screening of plasmid DNA by direct sequencing from bacterial colonies"

As with the previous method, this protocol uses the colony directly as a source of plasmid DNA for the sequencing reaction. Colonies of 1.5mm in diameter were picked with a flamed plastic tip and were resuspended in 3 μL of sequence dilution buffer (200 mM Tris pH9.0, 5 mM $MgCl_2$), 20 pmol (4 μL) of M13 forward primer (Life Technologies) was added, and the sample was briefly vortexed and centrifuged. The samples were heated at 96°C for 10 min in the thermal cycler. The tube contents were then spun down and 0.5 μL of BigDye Terminator Ready Reaction mix (Applied Biosystems) was added. The sample was quickly centrifuged again and returned to the thermal cycler (GeneAmp PCR System 2400 Applied Biosystems) for the cycle sequencing composed of 25 cycles at: 96°C for 10 s; 40°C for 25 s; and 60°C for 4 min.

6.3 Direct colony sequencing III: Zhang *et al.* (1999): "Multiplexed Automated DNA Sequencing Directly from Single Bacterial Colonies"

A colony of approximately 1.5mm in diameter was picked up using a flamed plastic tip, resuspended in 15 μL of deionized water and heated at 96°C for 11 min in a Progene thermal cycler (Mandel Scientific). After heating, the tube contents were spun down and 6 μL of BigDye Terminator Ready Reaction mix (Applied Biosystems), 2.5 μL of BSA (2 mg/mL) and 3.2 pmol (0.64 μL) of M13 forward primer (Life Technologies)

were added to the sample. The PCR sequencing cycling conditions used were: 96°C for 3 min followed by 40 cycles of: 96°C for 10 s; 50°C for 30 s; and 60°C for 2.5 to 4 min.

6.4 PCR-based sequencing I: Silva *et al.* (2001): "PCR template preparation for capillary DNA sequencing"

Amplification PCR was done on a fresh overnight culture, and the PCR fragments obtained from the amplification PCR were used as the sequencing DNA template. The amplification PCR protocol mentioned in the article was done as follows: 1 µL of 10× diluted fresh culture was added to amplification reactions containing 1×PCR buffer (Amersham Pharmacia Biotech Inc.), 125 µM dNTPs (Boehringer Mannheim), 1.5 mM $MgCl_2$ (Boehringer Mannheim), 1 pmol of each M13 reverse and forward primer (Life Technologies), and 1U Taq DNA polymerase (Amersham Pharmacia Biotech Inc.) in a final volume of 15 µL, adjusted with deionized water. However, in our hands, this PCR amplification protocol was not very efficient. We therefore used the PCR protocol described below instead. 1 µL of culture (non-diluted or 10× diluted) was added to the following: 0.5 mM of dNTPs (Boehringer Mannheim), 25 pmol of M13 forward primer (Life Technologies), 25 pmol of M13 reverse primer (Life Technologies), 5 µL of 10× PCR buffer (Amersham Pharmacia Biotech Inc.), 1 mM of $MgCl_2$ (Boehringer Mannheim), 0.5 µL of Taq DNA polymerase (50 000 U/mL, Amersham Pharmacia Biotech Inc.) for a total volume of 50 µL (adjusted with deionized water). The amplification PCR was carried out on a Mandel Scientific Progene thermal cycler and the cycles used were: 95°C for 4 min followed by 35 cycles of: 95°C for 45 s; 55°C for 45 s; and 72°C for 1 min, followed by 5 min at 72°C. Following the amplification PCR, 1 µL

of the PCR product was used for the sequencing reaction. 1 μL of PCR product was mixed with 6 μL of BigDye Terminator Ready Reaction mix (Applied Biosystems), 2.5 pmol of M13 forward primer (Life Technologies) and the volume was adjusted to 20 μL with deionized water. The sequencing PCR reaction was carried out at 95°C for 2 min, followed by 35 cycles at 95°C for 10 s, 50°C for 15 s, and 60°C for 1 min.

6.5 PCR-based sequencing II: Smith *et al.* (2000): "PCR-based setup for high-throughput cDNA library sequencing on the ABI 3700 automated DNA Sequencer"

10 μL of overnight culture was centrifuged to pellet the cells. The supernatant was removed and the pellet was resuspended in 50 μL of 10 mM Tris (pH8.0). After vortexing, 1 μL of the cell solution was transferred to a PCR tube containing 10 μL of PCR mixture which included: 0.1 mM each dNTPs (Boehringer Mannheim), 0.2U Taq DNA polymerase (Amersham Pharmacia Biotech Inc.), 1.1 μL of 10× PCR buffer (Amersham Pharmacia Biotech Inc.) and 0.5 μM each M13 forward and reverse primers (Life Technologies). However, we also modified this PCR amplification protocol as follows in order to generate better PCR products. 1 μL of the resuspended cells were added to the following: 5 μL of dNTPs (5mM) (Boehringer Mannheim), 5 μL of M13 forward primer (5 μM) (Life Technologies), 5 μL of M13 reverse primer (5 μM) (Life Technologies), 5 μL of 10× PCR buffer (Amersham Pharmacia Biotech Inc.), 2 μL of $MgCl_2$ (Boehringer Mannheim), 0.5 μL of Taq DNA polymerase (50 000 U/mL, Amersham Pharmacia Biotech Inc.) for a total volume of 50 μL (adjusted with deionized water). The amplification was performed using 30 cycles of: 94°C for 30 s, 60°C for 30 s, 72°C for 2 min. The PCR product was centrifuged to pellet bacterial debris and 2.5 μL

of the supernatant were transferred to a tube containing 7.5 µL of water. 20 µL of ethanol 100% were added and the samples were centrifuged to precipitate the DNA. The samples were dried down and resuspended in 0.5 µL of BigDye Terminator Ready Reaction mix (Applied Biosystems), 1.5 µL of sequencing buffer (Tris 200 mM pH9.0, 5 mM $MgCl_2$) and 3 µL of M13 forward primer (1 pM) (Life Technologies). The sequencing PCR was performed using of 25 cycles at: 96°C for 30 s; 50°C for 1 min; 60°C for 4 min.

6.6 PCR-based sequencing III: Truett *et al.* (2000): "Preparation for PCR-quality mouse genomic DNA with Hot Sodium Hydroxide and Tris (HotSHOT)"

This method used colonies instead of a liquid culture, and it has an additional lysis step. 75 µL of alkaline lysis reagent (25 mM NaOH, 0.2 mM disodium EDTA, pH 12) were used to resuspend the colony, after which 75 µL of neutralizing reagent (40 mM Tris-HCl pH 5.0) was added. From this cell resuspension solution, 1.5 µL were used for the amplification PCR, and mixed with 8.5 µL of PCR mixture composed of: 200 nM of each M13 reverse and forward primers (Life Technologies), 3 mM $MgCl_2$ (Boehringer Mannheim), 0.2 mM each dNTPs (Boehringer Mannheim) and 0.12U Taq DNA polymerase (Amersham Pharmacia Biotech Inc.). Similarly to the two previous protocols, the PCR amplification protocol was modified and replaced by the following protocol. Moreover, we found that the colonies did not had to be lysed to obtain a good PCR product, so PCR was also performed directly on the colonies. 1.5 µL of lysed cell solution (or a colony) was added to the following: 5 µL of dNTPs (5 mM) (Boehringer Mannheim), 5 µL of M13 forward primer (5 µM) (Life Technologies), 5 µL of M13 reverse primer (5 µM) (Life Technologies), 5 µL of 10× PCR buffer (Amersham

Pharmacia Biotech Inc.), 2 µL of $MgCl_2$ (Boehringer Mannheim), 0.5 µL of Taq DNA polymerase (50 000 U/mL, Amersham Pharmacia Biotech Inc.) for a total volume of 50 µL (adjusted with deionized water). The PCR amplification reaction was performed using 40 cycles of 95°C for 40 s and 60°C for 45 s. As for the sequencing protocol, none was mentioned in this article. The protocol used was therefore the one (Silva *et al.* 2001) mentioned in section 6.4.

A comparison of the PCR protocols for the PCR-based methods and our modified PCR protocol is presented in Table 12.

Table 12. Concentrations of dNTPs, $MgCl_2$, Taq DNA polymerase and 10x PCR buffer in the PCR fragment based methods and our modified PCR method.

	Our method	Silva *et al.*	Smith *et al.*	Truett *et al.*
dNTPs	0.5 mM	0.125 mM	0.4 mM	0.2 mM
Extra $MgCl_2$ added	1 mM	1.5 mM	-	3 mM
Primers	0.5 μM	0.067 μM	0.5 μM	0.2 μM
Taq DNA polymerase	0.5 U/μL	0.067 U/μL	0.018 U/μL	0.012 U/μL
10x PCR buffer	1X	1X	1X	1X
Total volume of PCR reaction	50 μL	15 μL	11 μL	10 μ

CHAPTER 7. Results

The results of six different sequencing methods have been compared (Table 13). Each method is also discussed in terms of its reliability, and of improvements that could improve its efficiency.

7.1 Direct colony sequencing I (Sasho 2000)

This method is easy to perform, does not require any pre-heating or lysis of the cell, but has a rate of success of only 59%, and the average length of the sequence is 376 bases. Even though the article calls for big colonies of 2-2.5mm in diameter that require at least 48 hours to grow, sequencing performed on smaller colonies worked well too, as long as the colonies were well resuspended.

7.2 Direct colony sequencing II (Brown Gladden *et al.* 2000):

This method's lack of success (9%) is probably due to the very low amount of Big Dye recommended (0.5 µL) and the small volume of the sequencing reaction (only 5 to 8 µL). But since this is the only direct colony sequencing method that has such a low volume, the failure to sequence could have been due to an agar contamination when the colony is picked up from the plate. Since the volume is much smaller than for the Sasho (2000) and the Zhang *et al.* (1999) methods, the agar contamination would be more concentrated and that could explain the lack of positive results. To verify that possibility, colonies were grown on agarose plates (a purer form of agarose) and these colonies were sequenced with two different method, the Zhang *et al.* (1999) method as a control and the Brown Gladden *et al.* (2000) method: here again, the Zhang *et al.* (1999) method gave

Table 13. Success rate, quality, time and cost of six recently developed sequencing methods.

	Sasho	Brown Gladden *et al.*	Zhang *et al.*	Silva *et al.*	Smith *et al.*	Truett *et al.*
# of templates sequenced	49	53	34	60	19	32
# of successful sequencing	28	5	31	38	0	20
% of success	59	9%	91%	63%	0%	63%
Quality	376 ±92	308 ±102	458 ±89	>210	NA	>210
% over 450 bases	29%	20%	71%	NA	NA	NA
Time required	+	+	+	++	+++	++
Amount of BigDye	6 µL	0.5 µL	6 µL	6 µL	0.5 µL	6 µL
Pre-heating	-	+	+	-	-	-
BSA	-	-	+	-	-	-

NA : does not apply: These three PCR fragment based methods were only tested on a 210 bp-long amplified fragment.
Quality = length at which the base calling sofware calls two undetermined bases per 20 bases.

very good sequencing results whereas the Brown Gladden *et al.* (2000) method gave no positive results.

7.3 Direct colony sequencing III (Zhang *et al.* 1999):

This method is the most reliable of all the six methods (see Table 13). Not only is it successful most of the time (91%), but it also gives the best reading length in terms of number of bases. In fact, its average of 458 bases of good sequence is about the maximum that can be expected from a short capillary on an ABI 310 sequencer. This superior performance in comparison to the other methods could be due to a combination of the pre-sequencing heating that can help lyse the cells, the addition of BSA (bovine serum albumin) and the large volume of the reaction that contains 6µL of Big Dye.

7.4 PCR-fragment sequencing I (Silva *et al.* 2001):

The results that were obtained with this method were good, but they were mostly obtained with a modified PCR method to prepare the PCR product. Our modified PCR protocol gave better PCR product with more intense signal than the protocol suggested in the original article. In Figure 8, the PCR product from lanes 1 to 6 is produced according to the Silva *et al.* (2001) method, whereas the more intense, therefore more concentrated PCR product in lanes 7 to 13 is produced with our modified PCR protocol (see section 6.4). The sequencing thus works best using our modified PCR protocol to generate the PCR template since that template is generated more often and is more concentrated.

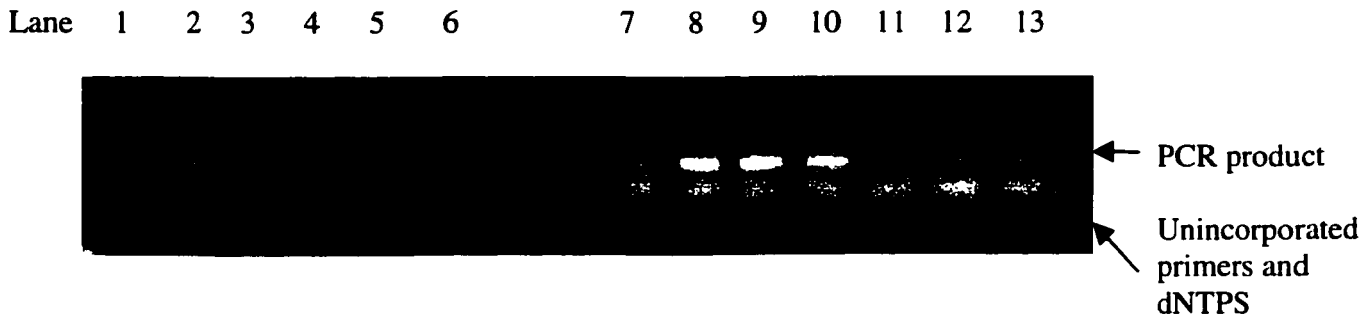


Figure 8. PCR products generated by two protocols, the Silva *et al.* (2001) method (lanes 1-6) and our modified PCR method (lanes 7-13). The upper band is the PCR product and the lower band is the unincorporated dNTPs and primers.

7.5 PCR-fragment sequencing II (Smith *et al.* 2000):

This method was not successful in producing any positive sequencing results, and it is the most fastidious one. The lack of results could be due to the very low amount of Big Dye used in the protocol (0.5 μL), similar to the Brown Gladden *et al.* (2000) direct colony sequencing method. Furthermore, the PCR protocol to generate PCR product that will serve as template for the sequencing encounters the same problem as does the Silva *et al.* (2001) method mentioned above; the protocol fails to generate strong PCR signal (therefore concentrated PCR products) and our modified PCR protocol was used to replace it. Figure 9 shows a good PCR signal in lanes 1 to 5 where the modified PCR protocol was used, while lanes 6 to 14 do not contain PCR product (generated with the Smith *et al.* 2000 method) strong enough to be visible on an agarose gel.

7.6 PCR-fragment sequencing III (Truett *et al.* 2000):

This method, which was originally designed to do PCR on mouse tissues, can also be used on bacterial colonies. It uses an alkaline lysis followed by a neutralization before the PCR is performed. Here again, our modified PCR protocol gives better PCR products as shown in Figure 10. The first five lanes illustrate PCR products obtained with our modified PCR protocol, while lanes 6 to 10 do not show any PCR product with the Truett *et al.* (2000) original method. Additionally, it has been found that the alkaline lysis steps and neutralization are not necessary to the success of the subsequent PCR. As shown in Figure 12, the alkaline lysis and neutralization steps do not increase the amount of PCR products, but mostly releases large genomic DNA fragments that are visible in the wells of the gel (lanes 6 to 10). The PCR product produced without alkaline lysis and

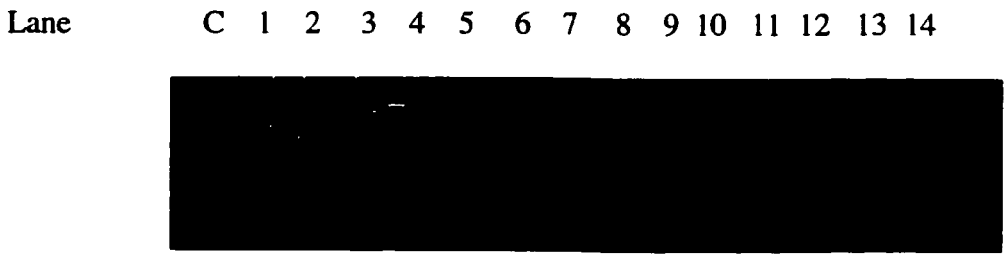


Figure 9. PCR products from two protocols, our modified PCR method (lanes 1-5) and the Smith *et al.* (2000) method (lanes 6-14). The upper band is the PCR product and the lower band is the unincorporated dNTPs and primers. C: negative PCR control

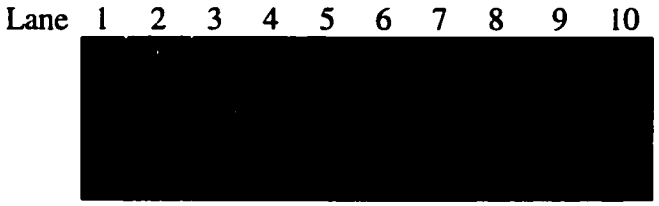


Figure 10 Comparison between two PCR methods, our modified PCR method (lanes 1-5) and the Truett *et al.* (2000) (lanes 6-10).

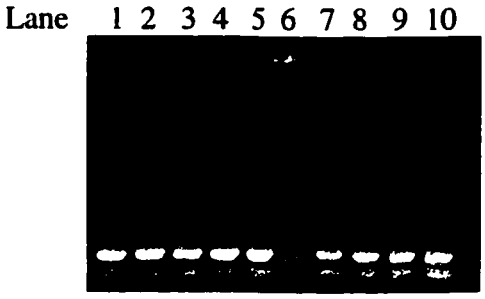


Figure 11. PCR products generated from an unlysed colony (without alkaline lysis) (lanes 1-5) and a lysed colony solution (with alkaline lysis) (lanes 6-10).

neutralization steps in lanes 1 to 5 are equally intense as those produced following an alkaline lysis.

7.7 Comparison of the six sequencing methods

The comparison of the six sequencing methods is summarized in Table 13 and clearly indicates that the best method regarding success rate or quality of sequence is the Zhang *et al.* (1999) method. For the direct colony sequencing methods, three factors seem to play an important role in these success rates: the high amount of Big Dye, pre-heating of the colony (prior to sequencing) and the addition of BSA. The most successful of these methods is the Zhang *et al.* (1999) method, and it does have all of these characteristics. Although the other two methods (Sasho 2000 and Brown Gladden *et al.* 2000) lack two of these items, the differences in success rate between these two indicates that the amount of Big Dye plays a more important role in favoring sequencing than pre-heating does. We have also demonstrated that our PCR protocol (for the PCR-based sequencing methods) is more adequate at providing reliable PCR amplification than the ones mentioned in the articles (Silva *et al.* 2001, Smith *et al.* 2000, Truett *et al.* 2000). In these protocols, the PCR reaction volume was small, in comparison to the one used in our protocol. It seems that for a small PCR reaction volume, the PCR does not work as well or may need more adjustment and fine-tuning than if using a larger reaction volume. All the PCR-fragment methods were not assessed for sequencing length since the PCR product that was used for these experiments was only 210 bases long. Furthermore, all these methods were successful at sequencing the full PCR product 100% of the time (when the PCR-sequencing reaction was successful). Some attempts were done at sequencing larger PCR

products, such as a 800 base PCR product and a 3kb PCR product. The success of these sequencing reactions seem to be related to the ease of producing a PCR product more than the sequencing itself. When sequencing the 800 base PCR product, the percentage of success was similar to that of the short PCR product (210 base long) and the read length did extend to around 400 bases (data not shown), which is what we would expect from a sequencing reaction run on a short capillary in an ABI 310 sequencer. Similar results were observed for the 3kb PCR product, though the PCR product was much more difficult to obtain due to its large size (data not shown).

CHAPTER 8. Discussion

8.1 Direct colony sequencing I, II and III

Out of the three methods based on direct colony sequencing, the Zhang *et al.* (1999) method was the most successful. All three methods have comparable aspects, such as time required, or sequencing PCR protocol. Although all methods call for different size colonies (note that the Zhang *et al.* (1999) does not specify any size), it has been observed that size is not particularly important, and that smaller colonies (around 1mm diameter) can be used for sequencing as long as they are well resuspended (data not shown). Although quite similar overall, these three methods do have some differences, and the latter are probably accountable for the large differences in the results regarding their frequency of success, quality of sequence and constancy in length of sequence. The parameter that seems to be crucial to the success of sequencing is the amount of BigDye Terminator Reaction mix (Applied Biosystems) used in the reaction. Since the BigDye Terminator Reaction mix contains most of the essential elements for the sequencing reactions to be successful, and that the manufacturer's instructions calls for 8 μL of BigDye Terminator Reaction mix, it is very probable that the use of only 0.5 μL of BigDye Terminator Reaction mix in the Brown Gladden *et al.* (2000) protocol is responsible for its lack of success, in comparison to the 6 μL used in the other two. Moreover, the most successful method (Zhang *et al.*, 1999) has an additional pre-heating step to lyse the cell, therefore releasing the plasmid DNA. This additional lysing step could explain part of the differences in success rate between the Sasho (2000) and the Zhang *et al.* (1999) methods. Finally, the Zhang *et al.* (1999) protocol also includes

bovine albumine serum (BSA) in its protocol, which likely helps to stabilize the DNA polymerase. It is surprising that this method works so well because the cells are simply lysed in water and the cell debris are not removed from the reaction. It might be possible to achieve a success rate close to 100% if some of the cell debris were removed from the reaction.

8.2 PCR-based sequencing I, II and III

The three PCR-fragment sequencing methods compared in these experiments have been adapted and modified to obtain comparable results. The Silva *et al.* (2001), Smith *et al.* (2000) and Truett *et al.* (2000) protocols all include a first amplification PCR reaction to obtain a PCR fragment containing adequate amount of DNA to perform sequencing. For all three methods, the PCR reaction done to obtain the DNA has been modified because the DNA amplification results that were obtained were not very good (see Figures 8-9-10). Without any good DNA fragment amplification, the second step, i.e. the sequencing, was not possible. Therefore, our modified PCR protocol mentioned in sections 6.4, 6.5, 6.6 is the one that was used for all the reactions. Nevertheless, the PCR conditions (time and temperature) were kept as described in each article. These lack of PCR amplification could be due to the small volume reaction described, or the wrong proportion of each components of the reaction (not enough dNTPs, not enough primers, lack of $MgCl_2$). The Silva *et al.* (2001) and the Truett *et al.* (2000) methods were equally successful, even though they differ slightly in the amplification PCR protocol. The Silva *et al.* (2001) method amplifies DNA from a liquid culture, whereas the Truett *et al.* (2000) method amplifies DNA directly from the colony, making it a shorter protocol

since the extra overnight growth for the liquid culture is not required. However, their similar results (see Table 13) are probably due to the fact that the sequencing PCR steps are identical (note that the Truett *et al.* (2000) protocol did not have a sequencing PCR step (it was not designed for this purpose); it was thus decided that the sequencing PCR protocol of the Silva *et al.* (2001) article would be used). As for the Smith *et al.* (2000) method, it is the only one that failed to produce any positive sequencing results, even though it was possible to amplify DNA using the modified PCR method. This lack of success is likely due to the very little amount of BigDye Terminator Reaction mix used in the sequencing reaction, i.e., only 0.5 µL. In addition to the lack of success of this method, it is the most labor-intensive method that we tested. Finally, the Truett *et al.* (2000) protocol suggests an alkaline lysis step prior to the amplification PCR, but our results have shown that it is not necessary. Although this alkaline lysis step is not detrimental to the reaction, it only serves to release more genomic DNA from the bacteria, which can perhaps be counter-productive in the sequencing PCR.

8.3 Comparison of all six methods

Overall, the Zhang *et al.* (1999) method is the one that provides the best results in terms of success and quality of sequence. It can sequence directly from a colony, making it faster since an overnight culture is not necessary, and it does not require a large size colony which can take at least 48 hours to grow. PCR-fragment based sequencing does have some advantages since it is possible to sequence DNA amplified from sources other than colonies or culture, like cDNA, but for bulk template sequencing, the Zhang *et al.* (1999) is without any doubt the most effective one.

8.4 Conclusion

The best method amongst those evaluated in this project is definitively the Zhang *et al.* (1999) method, giving its very good results of percentage of success and quality of sequence. This new method could be very valuable to many laboratories where a large amount of sequencing is done and where mini-preps can become expensive and labor intensive. It is a great screening method for clones, and once the clones of interest have been found through this fast and reliable sequencing protocol, the clones of interest can be further grown and purified if needed. Furthermore, we have improved the PCR protocol of the three PCR-based methods. Our protocol has a larger reaction volume, which seems to facilitate PCR amplification, whether it is done on a culture or a colony. We also showed that the alkaline lysis step of the Truett *et al.* (2000) method is not necessary to produce PCR fragments for sequencing.

REFERENCES

Birnboim H.C., Doly J. 1979. A rapid alkaline extraction procedure for screening recombinant plasmid DNA. *Nucleic Acids Research* 7: 1513-1523.

Brown Gladden J.G., Ray R., Gietz R. D., Mowat M.R.A. 2000. Rapid Screening of Plasmid DNA by Direct Sequencing from Bacterial Colonies. *BioTechniques* 29:436-438.

Cao Tin M., Ing Yao H.1991. Gel electrophoresis system for the micropreparation of plasmid DNA. *Analytical Biochemistry* 192:251-253.

Cole K.D. 1998. Preparative separation of plasmid and bacterial artificial chromosome DNA by density gradient electrophoresis in the presence of linear polymers. *Electrophoresis* 19:3062-3068.

Cole K.D. 1999. Reversible Gels for Electrophoresis and Isolation of DNA. *BioTechniques* 26:748-756.

DaRue P.A., Kinda H.K., Edward N.L. 2000. Separation and Purification of Plasmid Mixtures by Continuous Elution Electrophoresis. *BioTechniques* 29:1204-1206.

Davies M.J., Taylor J.I., Sachsinger N., Bruce I.J. 1998. Isolation of plasmid DNA using magnetite as a solid-phase adsorbent. *Analytical Biochemistry* 262:92-94.

Desruisseaux C., Slater G.W., Douin G. 1998. The gel edge electric field gradients in denaturing polyacrylamide gel electrophoresis. *Electrophoresis* 19 : 627-634.

Ferreira G.N., Cabral J.M., Prazeres D.M. 1998. Purification of supercoiled plasmid DNA using chromatographic processes. *Journal of Molecular Recognition* 11:250-251.

Ferreira G.N., Monteiro G.A., Prazeres D.M. and Cabral J.M. 2000. Downstream processing of plasmid DNA for gene therapy and DNA vaccine applications. *Tibtech* 18 : 380-388.

Holmes D.S. and Quigley M. 1981. A rapid boiling method for the preparation of bacterial plasmids. *Analytical Biochemistry* 114: 193-197.

Itoh M., Carninci P., Nagaoka S., Sasaki N., Okazaki Y., Ohsumi T., Muramatsu M. and Hayashizaki Y. 1997. Simple and rapid preparation of plasmid template by a filtration method using microtiter filter plates. *Nucleic Acids Research* 25 :1315-1316.

Itoh M., Kitsunai T., Akiyama J., Shibata K., Izawa M., Kawai J, Tomaru Y., Carninci P, Shibata Y., Ozawa Y., Muramatsu M., Okazaki Y, Hayashizaki Y. 1999. Automated Filtration-Based High-Throughput Plasmid Preparation System. *Genome Research* 9:463-470.

Levy M.S., O'Kennedy R.D., Ayazi-Shamlou P., and Dunnill P. 2000. Biochemical engineering approches to the challenges of producing pure plasmid DNA. *Tibtech* 18:296-304.

McNally M.T. 1999. Rapid, economical filter alternative to chromosomal DNA centrifugation in the alkaline lysis plasmid protocol. *BioTechniques* 27:68-71.

Prazeres D.M., Schluep T., Cooney C. 1998. Preparative purification of supercoiled plasmid DNA using anion-exchange chromatography. *Journal of Chromatography* 8:31-45.

Prazeres D.M., Ferreira G.N.M., Monteiro G.A., Cooney C.L. and Cabral J.M.S. 1999. Large-scale production of pharmaceutical –grade plasmid DNA for gene therapy : problems and bottlenecks. *TibTech* 17 : 169-174.

QIAprep. Principle Qiagen Kit. 11-37.

Sambrook J., Fritsch E. F., Maniatis T. 1989. Extraction and purification of plasmid DNA. Molecular Cloning: a Laboratory Manual 1.21-1.43.

Chikako Sasho 2000. Direct colony sequencing of plasmid DNA by dye terminator methods. *Biotechniques* 29 : 442-444.

Schluep T., Cooney C.L.1998. Purification of plasmids by triplex affinity interaction. *Nucleic Acids Research* 26: 4524-4528.

Silva W.A., Jr., Costa M.C.R., Valente V., de Freitas Sousa J., Paço-Larson M.L., Espreafico E.M., Camrgo S.S., Monteiro E., de Jesus Holanda A., Zago M.A., Simpson A.J.G. and Neto E.D. 2001. PCR Template Preparation for Capillary DNA Sequencing. *BioTechniques* 30:537-542.

Skowronski E.W., Armstrong N., Andersen G., Macht M. and McCready P.M. 2000. Magnetic, Microplate-Format Plasmid Isolation Protocol for High-Yield, Sequencing-Grade DNA. *BioTechniques* 29:786-792.

Slater G.W., Efcavitch J.W., Drouin G., Mayer P., Rousseau J., Zhou H.Y., Chiesa C., Ruhfel R., O'Neill R. 2000. Electrophoretic nucleic acid purification method. United States Patents 6 146 511.

Smith T.P.L., Godtel R.A., Lee R.T. 2000. PCR-Based Setup for High-Throughput cDNA Library Sequencing on the ABI3700™ Automated DNA Sequencer. *BioTechniques* 29:698-700.

Truett G.E., Heeger P., Mynatt R.L., Truell A.A., Walker J.A. and Warman M,L. 2000. Preparation of PCR-quality Mouse Genomic DNA with Hot Sodium Hydroxide and Tris (HotSHOT). *BioTechniques* 29:52-54.

Zhang Y., Tan H. and Yeung E.S. 1999. Multiplexed Automated DNA Sequencing Directly from Single Bacterial Colonies. *Analytical Chemistry* 71: 5018-5025.