# Clonal Selection and Characterization of Epigenetic Variation in *Pichia Pastoris*

by

Vasiliki Panagiotou

B.S., Chemical Engineering (2006)

Aristotle University of Thessaloniki, Greece

Submitted to the Department of Chemical Engineering

in Partial Fulfillment of the Requirements for the Degree of

Master of Science in Chemical Engineering

at the

Massachusetts Institute of Technology

May 2010

[June 2010]





Signature of Author .....

Department of Chemical Engineering

May 18, 2010

Certified by .....

J. Christopher Love

Associate Professor of Chemical Engineering

Thesis supervisor

Accepted by .....

William M. Deen

Chairman, Department Committee on Graduate Students

# Acknowledgments

As this step of my graduate career comes to an end, it is important to recognize and thank those people whose help was essential in completing this work and thesis.

First I would like to thank my advisor, Professor Christopher Love, for giving me the great opportunity to work with him. His support not only in research but also in non research related subjects has been very useful to me. Really, I have learned a lot by listening to all his comments and suggestions.

The Love Lab is a great environment to work in. I extend my thanks to all members of the Lab for being good friends and giving helpful advice, especially in my beginning. I would especially like to thank Kerry Love not only for being an amazing help in my research but also for being a true friend who would find a way to encourage me when nothing else could feel better.

Coming to this point I know I could not have made it without my family's support and love. My parents were always there to lighten my spirits and to encourage me to go on. A phone call at home, even though far away, could always make things better and give me strength to continue.

Finally, I would like to thank all my friends here in Boston for the unforgettable moments we shared. I will take with me the best memories of all of them.

## Preface

Portions of this work were adapted from the following publication that was co-written by the author :

Love, K., Panagiotou, V., Jiang, B., Stadheim, T. & Love, J. Integrated single-cell analysis shows *Pichia pastoris* secretes protein stochastically. *Biotechnology and Bioengineering* 2010

Clonal Selection and Characterization of Epigenetic Variation in *Pichia Pastoris*

by

Vasiliki Panagiotou

Submitted to the Department of Chemical Engineering

on May 18, 2010 in Partial Fulfillment of the

Requirements for the Degree of Master of Science in

Chemical Engineering

Abstract

Recombinant proteins produced by different host organisms have been broadly used as therapeutics. Considering the demand for large quantities of protein drugs, methods are needed to increase reactor titers in a timely and cost-effective manner. We used random chemical mutagenesis to modify a wild-type strain of the heterologous protein production host *Pichia pastoris*, which resulted in overall improvement of the secretion rate of the mutated population. More than 4000 single-cells were simultaneously screened for high secretion of a human Fc fragment using microengraving and the top-producing clones were retrieved. Future characterization of these improved clones by transcript profiling should yield information about networks of genes central in heterologous protein secretion in the yeast *P. pastoris.*

Table of Contents

# 1 CHAPTER 1:

Introduction

## 1.1 Introduction

The production of large quantities of biotherapeutic agents has become a significant need today. Most of these biotherapies are proteins, such as antibodies and catalytic enzymes, produced in host organisms using various recombinant techniques. The ultimate goal of each method is the same. On the one hand it is important to assure a viable and functional therapeutic product, and on the other hand, to increase productivity by creating high producing cell lines.

Challenges arise throughout the production of biotherapeutics. Some challenges include the choice of host organism, the cost of production, and the need for timeliness in process development. To enhance productivity, high producing clonal cell lines must be developed. This task requires selection of the best producers from a polyclonal cell population, and further development of these clonal cell lines. The inherent variability and heterogeneity in resulting clonal lines remains a key hurdle to the acceleration of cell line development[1]. Thus, high-throughput cell sorting and single-cell techniques that enable the isolation of high producers among a heterogeneous population are in great demand[2].

The work in this thesis focuses on developing a high-producing cell line for monoclonal antibody production using the yeast *Pichia pastoris*. The heterogeneity of a *P. pastoris* population was verified using an innovative technology called microengraving. The main question in this work was whether or not the secretion efficiency of *P. pastoris* could be enhanced by introducing random mutagenesis. Repeated rounds of chemical mutagenesis were performed, and at the end of each round, the best producers were chosen based on the

measured rates of secretion single-cell using microengraving. These high producing cells then served as parental clones for subsequent rounds. This method of iterative mutagenesis and screening did prove several improved *P. pastoris* clonal lines.

## 1.2 The use of recombinant proteins as therapeutics

Proteins are the molecular workhorses of biology: they are involved in the catalysis of metabolic reactions, they are the structural components of biological assemblies, and they are responsible for inter- and intracellular interactions and cell signaling events. A deficiency of protein production in the human body leads to several diseases that can be treated by clinical administration of the missing protein from external sources. Unfortunately, it is not always easy to obtain human proteins from their natural sources. Therefore, recombinant technologies for the production of heterologous proteins using different host organisms are needed.

In the past few decades, many techniques have been developed to engineer different organisms to produce recombinant proteins. Such proteins used as therapeutics have changed the face of modern medicine, since they provide innovative and effective therapies for a variety of diseases. Nowadays, their immense commercial value is well known in pharmaceutical industries and numerous studies have focused on finding ways to increase their production in a cost effective and efficient way.

The successful production of a protein is dependent on and related to the host organism and the machinery for production that it uses[3]. The choice of host cells has a profound impact

on the product characteristics and its maximum attainable yields. The pharmacokinetic properties of the products, which also depend on the host organism, are dictated by protein folding and post-translational modifications that affect solubility, stability and biological activity. Finally, product safety is a key aspect, so the production host should not allow the propagation of any pathogenic agents.

Several different organisms can be used for protein production based on the efficiency and the cost of the culture techniques required for each. The first recombinant pharmaceutical to enter the market was in the early 1980's, when the FDA (Food and Drug Administration) approved the clinical use of recombinant human insulin from *Escherichia coli* for the treatment of diabetes[4]. Since then, other recombinant drugs followed. So far, 151 recombinant pharmaceuticals have been approved for human use from the Food and Drug Administration (FDA) and/or by the European Medicines Agency (EMEA). These are predominantly produced in mammalian cells. Among the 151 proteins mentioned above 45 (29.8%) are obtained in *E. coli*, 28 (18.5%) in *Saccharomyces cerevisiae*, 17 (11.2%) in hybridoma cells, 1 in transgenic goat milk, 1 in insect cells and 59 (39%) in mammalian cells (Figure 1)[5].

Many of the host organisms used to produce these proteins are microbial cells, either bacteria or yeast. Despite the difficulties that these microbial systems might pose (lack or unconventional post-translational modifications and proteolytic instability) they are powerful tools for protein production.

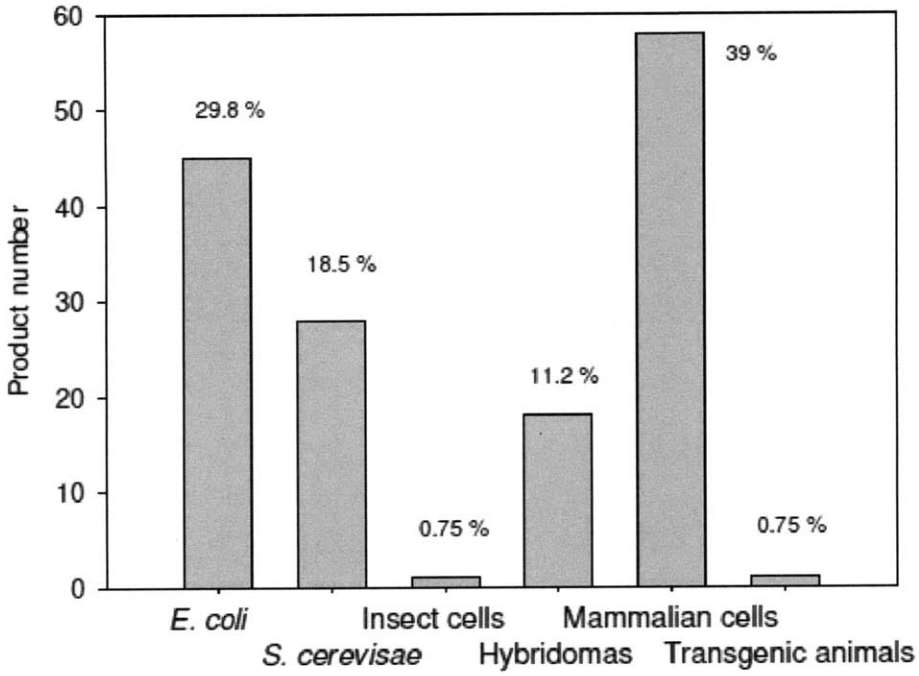


Figure 1 : Number of recombinant proteins approved as biopharmaceuticals in different production systems[5].

### 1.2.1 Monoclonal Antibodies used as therapeutics

Therapeutic monoclonal antibodies are a common class of therapeutics with applications in the fields of oncology, immunology and organ transplants. Their ability to enable the host to clear invading pathogens and their exquisite specificity for their targets have made antibodies an attractive choice of therapeutic for diseases where externally exposed membrane-bound or circulating proteins can be specifically targeted[6]. As such, antibody therapies such as Synagis, Herceptin, and Remicade have contributed to the treatment of infectious diseases, cancer and autoimmune diseases respectively[7]. Especially in the treatment of various cancers, mAbs have

proven to carry fewer side effects than the traditional cytotoxic drugs and have resulted in improved patient quality of life[8].

As of today, there are twenty two therapeutic monoclonal antibodies, or fragments, currently registered for marketing in the US (Table 1)[6]. The global antibody market is expected to grow from 17 billion US dollars in 2008 to more than 30 billion US dollars in 2010, with an annual rate of growth of 14%[9]. This statistic is based on the more than 500-antibody based therapeutics currently in development with more than 200 programs in clinical trials. Antibodies are considered to be among the most expensive of all drugs where for example the annual cost per cancer patient can reach $35,000. These high prices reflect the facts that antibodies are now marketed for chronic conditions and that their relatively potency is often poor, requiring high cumulative doses on the order of grams rather than milligrams[10].

Traditionally, the discovery of most therapeutic antibodies relies on the immunization of mice with the desired antigen. The cells that secrete antibodies against the injected antigen are fused with immortalized cells to allow growth for an extensive number of generations. Clones that produce the mAbs with affinity to the desired target are identified using classical affinity biochemistry methods[6]. A simple schematic of the procedure is shown in figure 2. The technique described above is the most commonly used. In addition, most of the production of the approved monoclonal antibodies relies on the cultivation of mammalian cells. This method is not only slow, considering the great demand of antibodies in market, but it also has limitations in some situations such as when generating a mAb against a toxin or a highly conserved antigen across species.

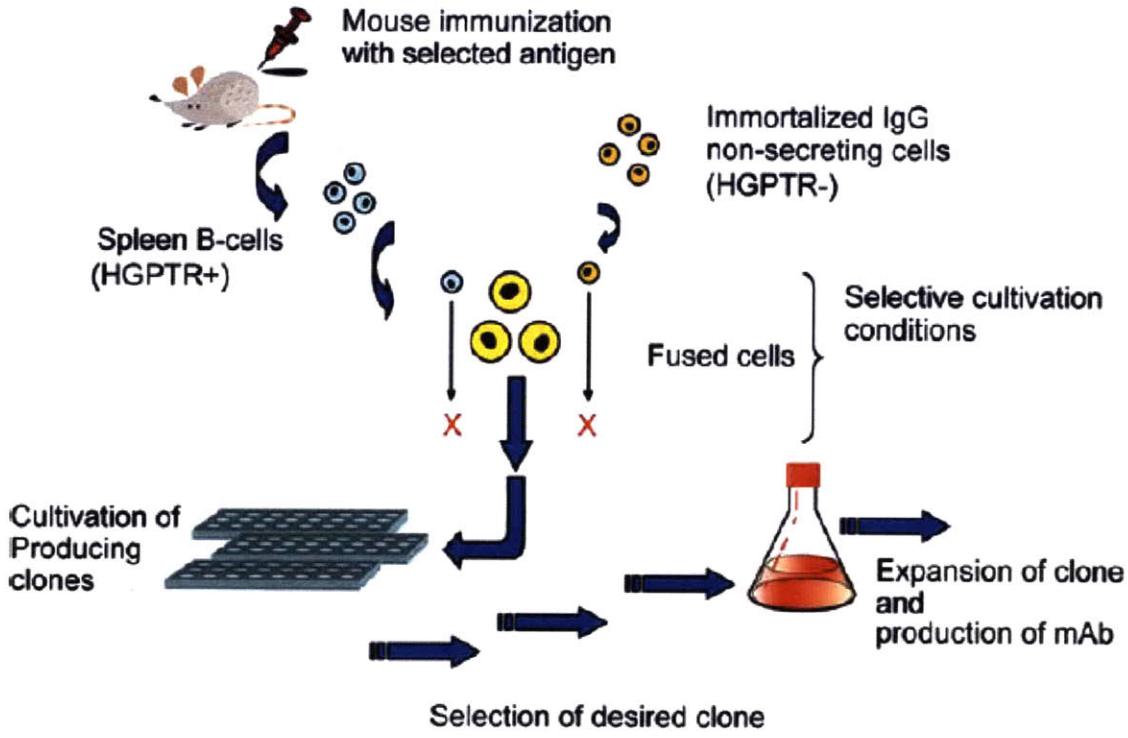


Figure 2 : Overview of the various steps leading to the production of monoclonal antibodies via the generation of fused hybridoma cells[6]

To overcome these limitations the generation of mAbs by recombinant technology was developed. Different production systems such as transgenic expression systems, *E. coli* and yeast, have been introduced and several antibodies have been produced from them. Transgenic production systems offer advantages in production cost, but are not expected to significantly alter development timelines or improve product attributes[11]. Microbial systems offer reduced timelines and acceptable expression levels, but lack post-translational modifications. Yeast culture has lower associated costs than mammalian cell cuture, but *N*-linked glycosylation from yeasts can be immunogenic[12]. Considering the above, there is a great challenge in choosing the right host organism for antibody production

## 1.3 Different host organisms used as production systems

As mentioned above, different organisms can be used as hosts for the recombinant protein production. The advantages and disadvantages of the most common hosts are discussed below.

### 1.3.1 Escherichia Coli as a host organism

The bacterium *E. coli* grows fast and its genetics are comparatively simple and easily-manipulated. Thus, it is the first choice microorganism for the production of recombinant proteins. So far, approved therapeutic protein-based products from *E. coli* are hormones (human insulin, calcitonic, parathyroid hormone, human growth hormone, glucagons, somatropin and insulin growth factor), interferones (alfa-1, alfa-2a, alfa-2b, gamma-1b), interleukins 11 and 2, light and heavy chains raised against vascular endothelial growth factor, tumor necrosis factor and several more[5]. Although it is widely used for several biotechnological applications, there are many disadvantages to using this organism. The most significant one is the necessity to renature most of the produced proteins, since in most cases, they are not properly folded. *E. coli* does not have endoplasmic reticulum (ER)[13]so recombinant proteins produced in this organism lack post translational modifications, that can play a crucial role in protein folding, stability and final biologic activity. Also, the frequencies with which codons appear in *E.coli* are different than human, and therefore, genes in *E. coli* may be inefficiently expressed leading to premature termination of protein synthesis.

### 1.3.2 CHO cells as host organisms

Mammalian cells, in general, are the dominant system for production of recombinant proteins for clinical applications, because of their capacity for proper protein folding, assembly and post-translational modification. CHO cells in particular, due to their adaptability to various culture conditions and plasticity in genetic alterations, dominate the domain of mass recombinant protein production. The first pharmaceutical produced in mammalian cells was tissue plasminogen activator and was synthesized using CHO cells, in 1987[14]. CHO cells have thereafter been used in pharmaceutical studies frequently, that they are regarded as the equivalent model of *E.coli* for mammalian cells. They possess most of the characteristics needed for an organism to be used as a host, and they have been demonstrated to be safe hosts for the synthesis of biologics. Specifically, CHO cells perform post-translational modifications and produce proteins with glycoforms that are both compatible with and bioactive in humans. Several other important factors have enabled CHO cells in industry, including ease of genetic manipulation and growth to high densities.

The challenges in working with CHO cells, or more generally, mammalian cells are dominated by the considerable degree of variability and variations among cultures (clonal heterogeneity) which leads to difficulties in scaling up systems for increased production yields. Another key issue that remains challenging in CHO cells is their expression instability; that is their ability to retain a constant level of protein production for long periods of culture. This trait is of particular importance for industrial companies who eventually hope to market the product.

Thus, although CHO cells are favorable mainly because of the post translational modifications they perform, there are important challenges with this type of cells that need to be taken into consideration.

### 1.3.3 *Saccharomyces cerevisiae* as a host organism

Yeasts are commonly used in industry for the production of recombinant proteins, mainly because of their ease of genetic manipulation, and the ability to grow in chemically defined medium in the absence of animal-derived growth factors (e.g., calf serum). Yeast cells can secrete large amounts of recombinant proteins and their fermentation is easily scaled-up. Researchers rely on yeasts for the production of recombinant proteins that cannot be obtained from *E. coli* because of folding problems or the requirement for glycosylation[15]. Currently, most of the FDA-approved therapeutic proteins in yeasts are derived from yeast type *S. cerevisiae* including hormones (insulin, insulin analogs, non glycosylated human growth hormone somatotropin, glucagon), vaccines (hepatitis B virus surface antigen) and virus-like particles of the major capsid protein of human papillomavirus type 6,11,16 and 18.

When searching for systems superior to prokaryotes for production of proteins, the baker's yeast *S. cerevisiae* is usually the initial choice. This organism secrets heterologous compounds via a multi-component secretory apparatus providing proper folding and post translational modifications like *N*- and *O*- glycosylation, phosphorylation and *N*-terminal acetylation[16]. Much is known about S. cerevisiae genetics, biochemistry, physiology and fermentation technologies to explain the vast use of this yeast type for recombinant

expression. Furthermore, *S. cerevisiae* is recognized by the FDA as an organism "generally regarded as safe".

Limitations and disadvantages have been encountered in the actual application of this organism as a host for protein production. These include instability of production strains , hyperglycosylation of many secreted glycoproteins with a consequent partial degradation[17] and relatively poor secretion efficiency[18]. *S. cerevisiae* is, therefore, not always the optimal choice of host organisms, especially when large-scale production is required, where fermentation needs require sophisticated equipment[19].

### 1.3.4 *Pichia pastoris* as a host organism

*Pichia Pastoris* is another yeast that has been widely used as a host organism to express recombinant proteins. It has proven useful for the expression of milligram-to-gram quantities of proteins in laboratories and in industry[5]. So far, *P. pastoris* has been used for the production of vaccines, antibodies, hormones, cytokines, receptors, ligands and many more. This host organism is suitable for proteins that form inclusion bodies in *E. Coli* and their expression level in mammalian cells is low[20]. As a yeast cell, it performs all the necessary post-translational modifications required for the functionality of a protein after its production. The advantages of this yeast over other expression systems are the rapid growth rate, the high level of productivity, the ease of genetic manipulation of its expression vectors, its capability for diverse posttranslational modifications, such as glycosylation and methylation, the strong promoter activity, and the ability to purify and engineer secreted proteins from medium without

harvesting the cells. *P. Pastoris* does not require special growth medium to grow to high cell densities[21] and compared with mammalian cells, it is genetically easier to manipulate[22]. For all these reasons, *P. pastoris* has become a popular and successful system for the production of heterologous proteins.

#### 1.3.4.1 The *P. pastoris* Expression System

Gene expression in *P. pastoris* is based on the fusion of heterologous gene sequences to strong methanol-inducible promoters and follows this algorithm : 1) insertion of the gene into the expression vector, 2) insertion of the vector into *P. pastoris*, 3) examination of potential strains for the expression of the foreign gene[23],[24]. For protein expression, a variety of host strains and expression vectors are used. A generalized diagram of an expression vector is shown in Figure 3.

Expression systems in *P. pastoris* may use different promoters, the most common of which is the methanol-induced alcohol oxidase (AOX1) that has been characterized and incorporated into a series of commercially available *P.pastoris* expression vectors[25]. Other types of promoters also used in *P. pastoris* are GAP, FLD1, PEX8, and YPT7. Among the above promoters mentioned, the glyceralhehyde 3-phosphate dehydrogenase (GAP) gene promoter provides constitutive expression on glucose, at a level comparable to that of the AOX promoter. A major advantage of using GAP as a promoter is that methanol is not required for induction and the culture does not need to be shifted from one carbon source to another. This characteristic makes GAP promoter strain growth and protein expression straightforward[20]. The

GAP promoter is less widely used, however, because the constitutive production of foreign proteins may have cytotoxic effects in the cells[26].

All *P. pastoris* expression strains are derived from NRRL-Y 11430 (Northern Regional Research Laboratories, Peoria, IL). Most of the strains have auxotrophic mutations that allow for selection of expression vectors containing the appropriate selectable marker gene upon transformation. Before the transformation, all strains grow on complex media and require extra nutrition for growth[23].

Finally, the genome of *P. pastoris* has been published[27],[28] and this report opens a new field in recombinant technology since the genome knowledge can help to highlight genes of interest in the secretory pathway or mprove existing industrial strains.

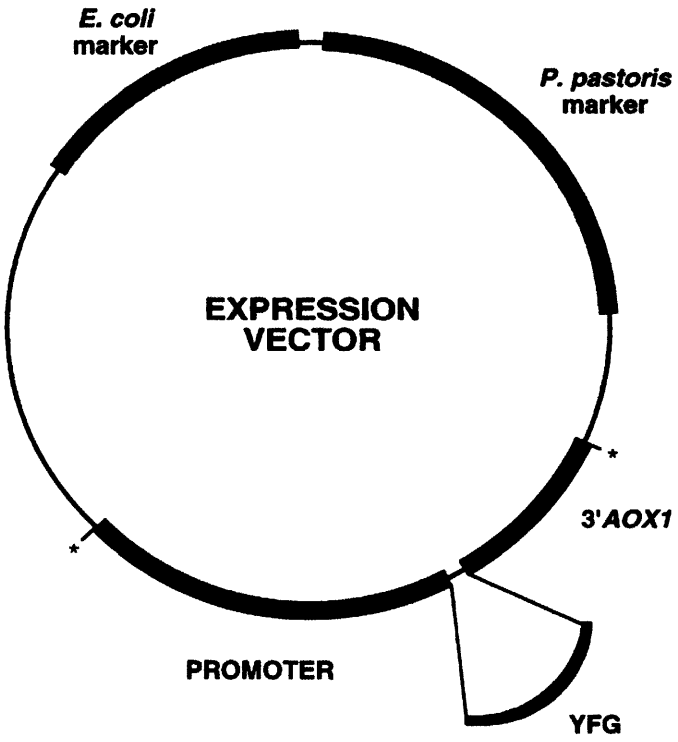


Figure 3: General diagram of a P. Pastoris expression vector[23]

### 1.3.4.2 Recombinant protein expression in *Pichia Pastoris*

Like other types of yeasts, *P. pastoris* needs sources of carbon and nitrogen to grow. The most common carbon sources are glucose and glycerol and nitrogen sources are peptone, yeast extract and yeast nitrogen. It is thought that medium composition matters in protein expression by influencing cell growth and viability[29]. *P. pastoris* can grow in minimal media to high cell densities and genetic stability of the recombinant protein is helped with integrated vectors.

A major advantage of *P. pastoris* over other systems, especially bacteria, is that it can perform post-translational modifications associated with higher eukaryotics. One of the most important and critical modifications in *P. pastoris* is protein glycosylation[20].

Glycosylation is the most extensive post-translational modification, and it affects the secretion, antigenicity and clearance of glycoproteins. It requires enzymatic process where saccharides are added to proteins to produce glycans. In the case of yeast, the addition of many mannose residues to the core oligosaccharide often compromises recombinant protein activity[30]. To be more specific, yeast utilize a conserved *N*-linked oligosaccharide biosynthetic pathway that involves the formation of a Glc3Man9GlcNAc2-PP-dolichol lipid-linked precursor, the glycan portion of which is transferred co-translationally in the endoplasmic reticulum (ER) to suitable Asn residues on nascent polypeptides[20]. In *P. pastoris*, the oligosaccharide chains attached to proteins are shorter as compared to *S. cerevisiae*; there are only 8-14 mannose residues in *P. pastoris* as opposed to *S. cerevisiae* where there are 40-150[31]. Another advantage of *P. pastoris* glycosylation is that the oligosaccharides secreted by *P. pastoris* do not have a

terminal α-1,3-linkages like *S. cerevisiae* and the site of glycosylation is Asn-X-Ser/Thr, like in mammals.

Many glycosylated proteins have been successfully expressed in *P. pastoris*, but protein-linked carbohydrate biosynthesis is different between this yeast and organisms such as human. One method to improve that process is by engineering *N*-glycans to make them human-like. In this case, the high mannose yeast glycans are converted *in vitro* following expression to mammalian type high mannose glycans[32]. Another method used combinatorial genetic libraries to humanize *N*-linked glycosylation, where the secretory pathway was reengineered in a way that nonhuman glycosylation is deleted. Combinatorial genetic libraries were constructed and added to *P. pastoris* to localize active alpha-1,2-mannosidase and human beta-1,2-*N*-acetylglucosaminyltransferase I in the secretory pathway[33].

## 1.4 Improvement of host properties

The need for therapeutic proteins to be used as drugs, many of which are antibodies, requires large doses (currently between 15-60 pg/cell/day[1]) and voluminous production quantities. For each therapeutic product, a cell line with sufficient production capability must be developed. Cell line development spans over a period of time, around 6 months depending on the cell type, and involves screening of high producing cells until the top producing clones are isolated. Such development of a cell line is challenging, not just because of the different cloning techniques that are required for different cell types, but mainly because screening

techniques are needed to identify the high producing clones from thousands of potential producers.

In order to make these "hyperproducers", the properties of the cells that are related to the production of proteins often need to be improved. The productivity of a recombinant system depends on several genetic and physiological factors, such as the codon usage of the expressed gene, the gene copy number, the efficient transcription by promoters, correct processing and folding in the endoplasmic reticulum and finally successful secretion out of the cell. It might also depend on other environmental conditions that affect the growth of the cells and the production of the protein. An often used approach to improve the production properties of the cells is to change their internal and/or external environment.

Modulating the external environment can include varying operating conditions, such as growth temperature or nutrient concentrations in growth media. Cultivation pH and oxygen supply also can enhance productivity. Medium composition is thought to influence heterologous protein expression in yeast by affecting cell growth and viability[34]. Especially for *P. pastoris* it has been shown that yeast extract, casamino acids, or EDTA may enhance protein accumulation[35]. Furthermore, low temperatures lead to reduced protease levels and increased yields of foreign gene products[36]. Finally, pH conditions affect proteolysis and stability of proteins. For *P. P*astoris an optimum pH is around 3-6, but this condition may vary depending on the protein produced[37]. Finding the optimum conditions for growth and production is indeed important, but perhaps less critical than obtaining a stable host that is consistent in protein production from batch-to-batch.

Intracellular changes to the genes related to protein production are affected either by mutagenesis or recombinant DNA (rDNA) approaches. Performing the latter, changes occur in the genome of the cells that ultimately have an effect on the expression system of each organism. When rDNA methods are used to increase productivity, different genes and secretory pathways are affected. There are numerous related studies reported in the literature, a few examples of which are:

i) *Insertion of gene of interest and screening for the high producer:* Searching for the best producer after inserting the gene with cloning techniques is the most common way to produce recombinant proteins. Many efforts have been successful in protein production. However, simply inserting a gene of interest into a vector and transforming a microbial host is no guarantee of a viable process[38].

ii) *Promoter engineering to enhance productivity:* In the case of *P. pastoris* this has been done by deletion and duplication of putative transcription factor-binding sites in the AOX promoter and resulted in improvement of yields and quality of the produced proteins, as well as in a tool-box to fine-tune gene expression[39]

iii) *Optimizing codon usage and increasing the gene dosage*: Improvements in secretion of recombinant proteins has been achieved by overexpression of genes in different hosts such as CHO cells[40], *E.coli*[41] and *P. p*astoris[42] where changes in the endoplasmic reticulum (ER) resident, homologous chaperone protein, protein disulfide isomerase (PDI) were able to increase the secretion of proteins in high copy clones.

iv) *Specific mutagenesis introduced by error-prone PCR*: In this case, the gene of interest is mutated by changes in one or more amino acids by error-prone PCR. This technique allows in substitution of amino acids that would ultimately affect the gene function. Successful changes that lead to high producing cells are identified with screening techniques.

A much faster way of changing the internal structure of a gene without molecular cloning, an often laborious process, is random mutagenesis. Mutations of different types can be induced in host cells. The two most common mutagens in the case of yeast cells are ethyl methanesulfonate (EMS) and ultraviolet (UV) light. Random mutagenesis can increase the frequency of mutation up to 100-fold per gene, with about 70% killing of cells and without a significant frequency of double mutants[43]. EMS and UV produce different spectra of mutants and only one of the two types is sufficient to generate enough number of mutants to study [34]. EMS produces random mutations in genetic material by nucleotide substitution and particularly by guanine alkylation. This treatment typically produces only point mutations and induces mutations at a rate of $5x10^{-4}$ [44] . The mutations with UV light occur as a repair in the DNA caused by the damage of the light. Such repair is called "error-prone"[45].

The libraries of variants achieved from either mutagenesis or rDNA methods must be further screened to identify the best producers. Different screening techniques and assays have been used[46]. In *P. Pastoris* for example, potential transformants for a particular gene of interest can be first screened on yeast extract-peptone-dextrose (YPD) plates with a selective marker to verify foreign gene insertion, and then the amount of protein secreted can be measured by

sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE), western blotting, or flow cytometry[47]. The challenge in each screening technique is the difficulty to isolate the high producers from a heterogeneous population consisting of high, medium and low producers. The presence of low and medium producing cells that usually dominate the culture makes screening and identification of high producers hard. Since the ultimate goal is the improvement and optimization of therapeutic protein production, it is important that screening methods are developed to select for high-producing cells within heterogeneous populations in a timely and cost effective manner.

## 1.5 Heterogeneity among cell populations

Recent advances in cell screening techniques have proven the heterogeneity among populations. Analysis of cells in culture has shown that with the exception of DNA content, all other cellular components are distributed over a wide range, showing a large amount of deviation in growth characteristics[48]. Indeed, the theory that all cells within a population are genetically and phenotypicaly identical is strongly disapproved[1]. Variability is an inherent characteristic and does not arise simply as a result of the metabolic burden imposed by recombinant protein expression[1]. The heterogeneities are caused by genetic variations in a culture, by the progression of cells through the cell cycle and by inhomogenous cell microenvironments[49].

Even with repeated rounds of cloning or in the presence of selective pressure an entirely homogeneous cell line does not arise. Indeed, an entirely homogeneous population is an

unrealistic prospect as seen in industrially important cell lines such as CHO and hybridoma cells[1]. The heterogeneity is a barrier for increased product yields and even more, a decline in production may be seen because of the presence of medium and low subpopulations. The main reason that causes the latter is the overgrowth of such subpopulations because high-producing cell metabolic resources are targeted towards protein production so they have to lower their growth rates[50].

Considering the above, an accurate and reliable cell sorting technique that provides with the distribution of parameters within the population and gives the opportunity to isolate the subpopulation with the desired properties is needed for an accurate study of cell populations.

## 1.6 Clonal selection of best producers

A key requirement for the production of therapeutic proteins in industrial cell culture is that the producing cell line is clonal; that is, derived from a single cell[51]. Single-cell analysis is becoming more critical as investigators are becoming aware of cellular heterogeneity, and as a consequence many techniques that help to identify and isolated single cells are emerging.

High-throughput screening techniques that identify single cells based on their production ability are desirable. Such techniques also should be reliable, objective, cost efficient and easy to use. Several such methods are currently available, a few of which are reviewed below.

### 1.6.1 Flow cytometry

Flow cytometry (FC) is the first single-cell analysis method that can describe the distribution of the cell properties within a population, allowing cell viability throughout the procedure[48]. It is one of the very few methods available not only to select for single cell but also to quantify the relative production of a specific compound. Compared to other single-cell tracking methods, it provides several advantages, such as higher accuracy and reproducibility, and significantly shorter analysis times[52]. In this method, the cells travel in a liquid stream, and as each single cell passes through the exciting light and the measuring optics, it sends out a number of size and structure signals as well as fluorescence signals which depend on the staining procedure that has been used. The information for each cell is stored independently and can be further analyzed.

One version of flow cytometry is the Fluorescence-Activated Cell Sorter (FACS), which allows simultaneous staining, analyzing, and then sorting of cells from small samples. In FACS, a wide variety of cell surface markers can be detected with fluorescently labeled monoclonal antibodies. So far, up to 20 different fluorescent colors, which are translated to 20 different cell parameters, can be measured simultaneously in a small sample to quantify gene expression[53]. The advantages of multiple fluorescence parameters measurements are not just the economy of reagents and labor, but also the geometrical increase of information obtained given a number of parameters. An example of multiparameter analysis is shown in Fig. 4 where 3 different fluorescence signals can be measured to identify 3 different subpopulations within the cell population examined.

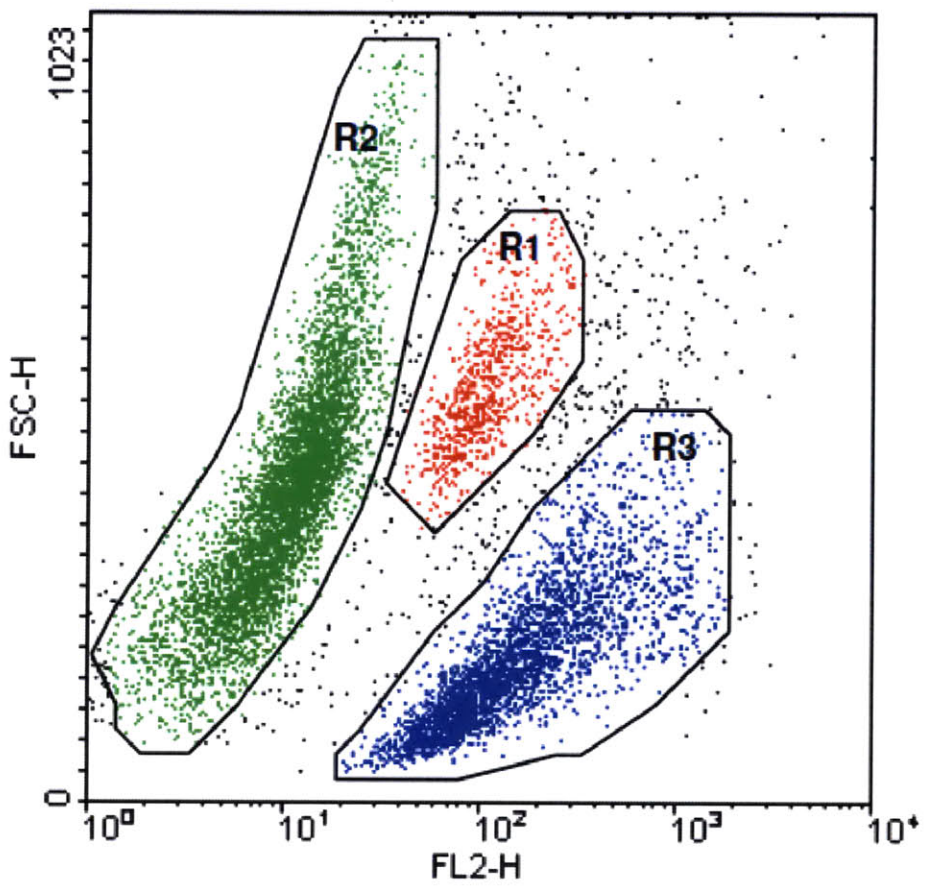


Figure 4 : Multiparameter analysis methods. By gating on different subpopulations, their properties with regard to several parameters can be determined[48].

With FC, the analysis of $10^4$ cells is a standard procedure[49]. The ability of this technique to sort the rare producers of interest within a population of millions of cells makes it favorable in biotechnology. Potential applications are widespread and very versatile, ranging from screening for specific features of different biomolecules, such as overproduction, to cell and protein engineering, where screening for the optimized cells is performed from a recombinant library. Flow cytometry has also been used to quantify the production of recombinant proteins in yeasts[54]. When compared to other methods, FC is superior for analyzing yeast because the cell wall is rigid, and this makes the quantitative cell lysis of protein problematic. A summary of

a few of the applications of flow cytometry reported in different cell types are shown in Table 2. Flow cytometry can also provide powerful results when used to follow a fermentation process. There are numerous studies in literature where cell viability is measured after high cell density cultivation, for example in *P. Pastoris*[1]. The results of such studies serve as great input in industry since there are not many good process analytical technologies available right now to gain info about products during fermentation.

Table 2: Overview of cell sorting applications in biotechnology

| General aim | Sorting target | Selected examples |
|---|---|---|
| Physiological research | Viability, vitality | bacteria, yeasts |
| Protein engineering | ligand binding | antibody surface display |
| | | peptide surface display |
| | enzyme engineering | intra- and extracellular enzymes |
| Cellular properties | cell hybridization, cloning | yeast hybridization, library cloning |
| | promoter trapping | bacteria |
| | robustness | acid tolerance |
| | process related properties | high cell density, low growth rate |
| Overproduction | product stained by immunofluorescence | protein |
| | Autofluorescence of product | alkaloids |
| | Unspecific staining | FITC/antibiotic production |

## 1.6.2 Gel microdrop technology (GMD)

Gel microdrop technology is a flow cytometric alternative to the ELISA-based cloning process, and it was developed to assay multiple parameters of individual cells simultaneously. The proteins secreted from a cell are captured in a gel matrix and quantified with a fluorescently labeled antibody. GMD has expanded the types of assays that can be used with

flow cytometry, since this method not only measures cell surface fluorescence but also can be used to monitor changes that take place in the extracellular region[52].

Gel microdrops are formed by emulsifying liquefied agarose containing a cell suspension in a non-aqueous medium. A specific capture antibody binds to the biotinylated matrix via a streptavidin "bridge". The cells trapped in the microdrop secret proteins that bind to the capture antibody and are then detected by another fluorescently labeled conjugated antibody or a specific antigen[48]. In such cases, high producers are identified by relative fluorescence intensity. The principle under which GMD technology performs is shown in Figure 5.

The GMD assay can be readily adapted to measure a variety of characteristics simultaneously, including viability, secretion and surface marker specificity. It was first used in bacteria[53], and since then, it has been used in fungal and mammalian systems[55] for applications such as separation of secreting to non-secreting hybridomas[56] or detection of high and low secreting populations of cells[57]. It has also been used in conjunction with FACS to analyze and enhance secretion in hybridomas[58]. The great advantage of this system is the restriction in product diffusion and high saturation. A minor draw-back, however, is that, to ensure single cell occupancy, low cell density is seeded, resulting in only 10-15% of beads containing single cells.

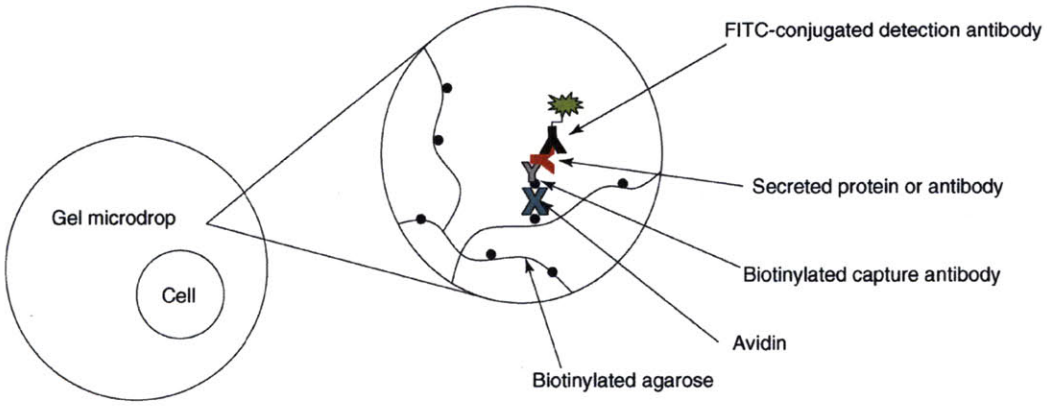


Figure 5 : Gel microdrop technology. Cells are encapsulated in a biotinylated agarose droplet; an avidin bridge links biotin (black circles) to a biotinylated antibody specific to the protein or antibody secreted by the cell, secreted protein is immobilized on this matrix and detected with an FITC-conjugated detection antibody [44].

### 1.6.3 Matrix-based secretion assays

This method is quite similar to GMD-type assays. Here, the secreted protein is immobilized on an artificial matrix on the cell surface. The cells are biotin labeled and are either tagged with avidinated capture antibody or via an avidin bridge to a biotinylated capture antibody. The latter case is more efficient since the avidin linker maximizes the binding capacity of the matrix, and biotinylated ligands (compared to avidinated ones) are more readily available on market[1]. The bound protein is labeled with a fluorescence tag so then cells with the desired properties can be identified. A schematic figure of this method is given in Figure 6. In this assay, diffusion of the secreted protein is prevented by incubation of cells in high-viscosity medium. This method has shown to decrease the duration of the selection process by almost 25%[1].

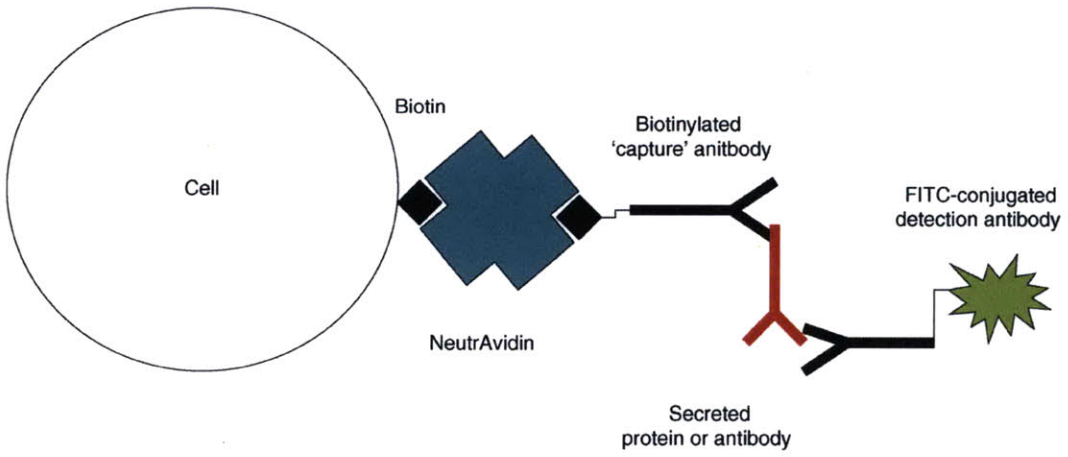


Figure 6 : Affinity capture surface display. Biotinylated cells are linked to a biotinylated 'capture antibody specific to the secreted protein or antibody via a NeutrAvidin bridge (blue cross), cells are allowed to secrete into a medium of high viscosity retaining secreted protein in the vicinity of the cell. Bound antibody is detected by a fluorochrome conjugated detection antibody[1].

### 1.6.4 Microfluidic Devices

Microfluidic systems are very popular among the single cell techniques used. They allow for the facile manipulation of cells in very small volumes, on the order of nano-liters, and have enormous potential for enabling measurements of secreted proteins from isolated single cells. Microfluidic systems typically allow four to five times reduction in the sample volumes as compared to traditional assays and as a result, a reciprocal increase in analyte concentration is obtained such that the expressed protein can be easily detected. There are many types of devices requiring slightly different measurement methods used in the microfluid area, the principle under which they function though remains the same.

Microfluidic chambers consist of channels and valves to control flow. These independent channels and microvalves create the fluidic network that is used to isolate cells

and their secreted proteins. Most devices are made of polydimethylsiloxane (PDMS), since it provides excellent mechanical and optical properties and has solved many of the miniaturization and automation challenges of microfluidic immunoassays[58]. The surface of the devices may be specially treated to change the chemistry and allow for protein binding, in many cases with capture antibodies. Proteins secreted from cells flowing in the microchambers are captured by these immobilized antibodies performing an enzyme linked immunosorbent assay (ELISA)-like fluorescence immunostack. The fluorescence signals from these microchambers quantify the captured antigens[59]. A schematic figure of the device is shown in figure 7.

The ease in use and the relative cheap fabrication of such devices makes them highly favorable. Also, the large amount of single cell events that can be measured, as well as the multiplexing capability to detect more than one antigen at a time, allows broad applicability of this technique. A key disadvantage is the fact that cells often cannot be retrieved for further measurements.

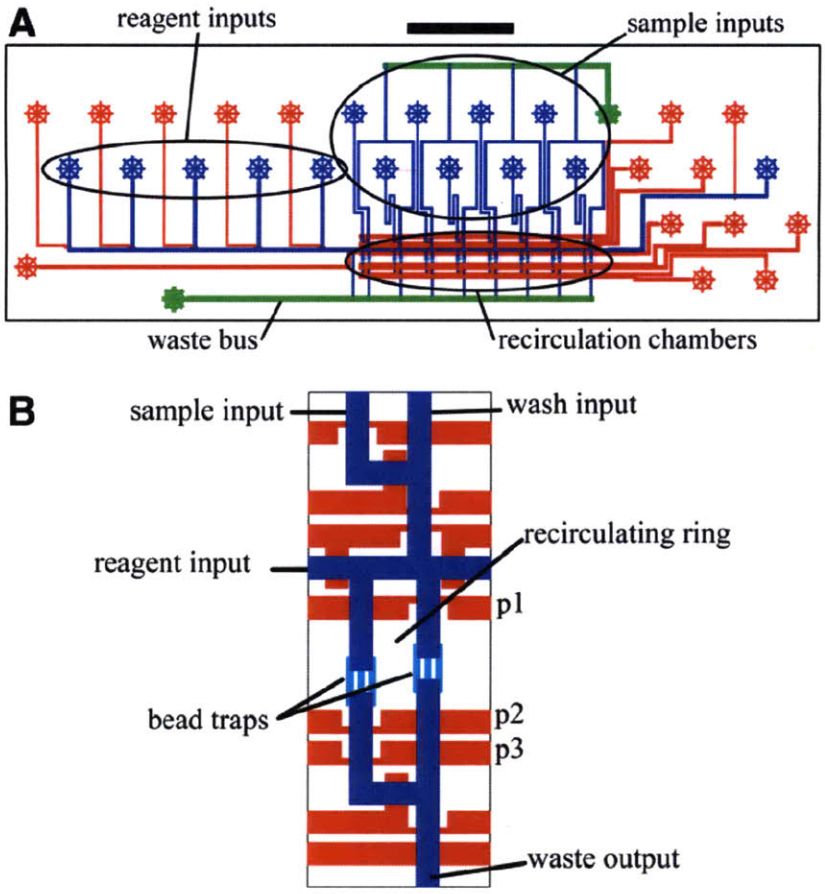


Figure 7 : Microfluidic Chip design. (A) Control channels are shown in red (23 lm height), and flow channels are shown in blue (13 lm height) or green (65 lm height). (B) Detail of one recirculating sample chamber. Control channels are shown in red (100 lm width), and flow channels are show in blue (13 lm height) or light blue (2.5 lm height). The three valves forming the recirculating peristaltic pump are labeled p1, p2, and p3.

### 1.6.5 Commercial high speed machines for single cell measurements

#### i) Genetix ClonePix FL technology [60]

This technology uses fluorescence imaging to enable screening of a heterogeneous population for the selection of the highest expressers for a particular receptor, or combination of receptors/cell surface proteins, as well as the isolation of single cells to create monoclonal populations, as opposed to flow

cytometry where only the enriched top 5% population can be isolated which is polyclonal. This technology has also been extended to the selection of colonies expressing transfected receptors with extracellular epitope tags encoded in their constructs. Again, expression is quantified using fluorescently conjugated antibody to that epitope tag for detection.

ClonePix FL technology combines the fluorescence detection with the clonal selection and has improved the timeline, labour costs and overall efficiency of selecting clonal cell lines based on surface protein expression. The number of colonies that can be screened are up to 3000.

ii) **Cyntellect Leap System**[61]

This technology platform uses an ultra high-speed F-theta optical scanner and real-time image analysis to phenotype cells. It also incorporates a high-speed targeting laser that allows single cell selection based on their phenotype. With this technology, cells are processed in situ, thereby expanding the types of cells that can be processed, increasing efficiency of processing, improving information content, and enabling novel cell-based experimentation. In such technology, either 96 or 384 well plates are imaged using brightfield and fluorescence. More than 1000 cells/sec of different types, including primary cells, are manipulated with high-speed in situ laser on two different wavelengths (355, 532nm) and are gated based on fluorescent or non-fluorescent phenotypic and functional criteria. The viability of the cells after manipulation is high.

The main disadvantage of both commercial technologies mentioned above is that they use a highly automated cell culture system which is an expensive investment and it might need to be changed when a new one enters the market. Also, Cyntellect LEAP systems works only with adherent cells.

## 1.7 Dissertation objectives

The Love Lab is concerned with the development of new methodologies to explore the heterogeneity present in populations of cells and to further characterize the dynamic biological responses of individual cells. Microengraving is an emerging method developed by the lab and it is based on soft lithography. The applications of this technique are tremendous, not just because of the ability to measure multiple characteristics of single cells, but also because of the large number of cells that are simultaneously measured that may subsequently be retrieved.

Using microengraving, the main objective of my work was to enhance antibody secretion and productivity of a yeast strain, through repeated rounds of mutagenesis. Chapter 2 describes experimental details of my experiments using microengraving to screen populations of mutant yeast. A deep characterization of yeast cell populations based on their heterogeneity is also provided. Cells retrieved following three rounds of mutagenesis yielded more secreted protein material than the original parent clone. A discussion of the potential epigenetic and genetic effects as well as the selection criteria in order to achieve better producers is provided in Chapter 3

# 2 Chapter 2 :

## Clonal selection and mutation in *Pichia Pastoris*

## 2.1 Introduction

This chapter presents the methodology used to enhance secretion in a cell population. The yeast *Pichia pastoris* was used as the host organism to secrete a human Fc fragment. P. pastoris was randomly mutagenized to introduce changes leading to an increase in productivity and protein secretion was measured using the microengraving technique. The clones secreting protein at the fastest rate were identified by microengraving and retrieved from the population using an automated cell-picking robot. Further characterization of the best producing clones was done by ELISA.

In the beginning of this chapter, the microengraving technology is introduced, including a reasoning of why it was selected as the best suitable technique for measuring single cell secretion and retrieving the cells of interest. Initially, we characterized the Fc-secreting strain of *P. pastoris* by measuring the secretion of Fc over time. The heterogeneity of the clonal, single-cell population is presented. Finally, a chemical mutagenesis methodology is presented where, after 3 rounds of random mutagenesis, an overall better secreting population, compared to the parent strain, was achieved. A full characterization of these better clones is also provided.

## 2.2 What is microengraving?

Microengraving is a soft lithographic technique that uses a dense array of microwells, 0.1-1 nl each, and a common glass slide to print protein microarrays, wherein each feature comprises the protein secretions from a single cell. This technique enables not only identification of those cells exhibiting a desirable secretion-based phenotype, such as secretion of an antigen-specific antibody, but also subsequent recovery of the single cells for clonal expansion[62]. After the microengraving process, cells remain viable in the microwells and the corresponding protein microarrays are interrogated in a manner similar to commercial microarrays of proteins or antibodies.

The arrays are fabricated by a combination of photolithography and replica molding of monolithic slabs of poly(dimethylsiloxane) (PDMS)[63]. PDMS is a biocompatible material, that is non-toxic, it is gas permeable and it is easily compressed to form a tight, but reversible, seal with a rigid substrate[64, 65]. The arrays comprise blocks of wells either 50 or 30μm in diameter and depth. Each well is separated from the other by 50 or 30 μm correspondingly. The dimensions of the array are 1" x 3" and it fits within the boundaries of a commercial microarray glass slide. A representation of the microwells is given in Figure 2. The surface of the PDMS slab is treated for 30s with an oxygen plasma and then immersed in PBS to increase wettability and to minimize adhesion of cells to wells.

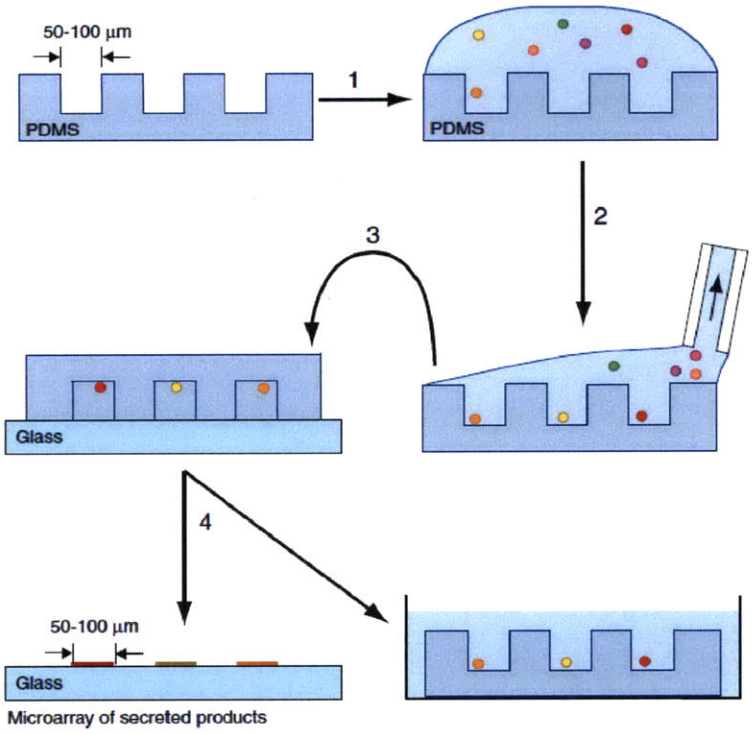


Figure 1 : Schematic diagram depicting method for preparation of engraved arrays of secreted products from a mixture of cells. (1) A suspension of cells is deposited onto an array of microwells fabricated by soft lithography. (2) The cells are allowed to settle into the wells and then the excess medium is removed by aspiration. (3) The dewetted array is placed in contact with a pretreated solid support, compressed lightly and incubated for a few hrs. (4) The microwells are removed from the solid support and placed in a reservoir of medium [1].


Figure 2 : Photograph of a PDMS microarray. Each microwell in 50µm in diameter with the entire array comprising ~85,000 features.

To prepare the microwells for engraving, cells are deposited on the surface of the PDMS at the appropriate dilution to maximize single-cell occupancy. The number of cells deposited into each well depends on the concentration of cells, the volume applied, the time allowed for settling, the size of the microwells and the size of the PDMS slab[62]. Once the cells are settled in the wells, they are sealed against a glass slide, where a capture antibody has been immobilized, and left to secrete proteins for an empirically determined time. Following protein secretion and deposition, the glass slide is removed and subsequently treated for further analysis. The cells remain in the wells, where they can survive for a few days when supplied with enough medium.

During microengraving, the antibodies secreted from the cells are captured on the glass slide and detected by a fluorescently labeled secondary antibody. The intensity of the fluorescence corresponds to the amount of protein secreted by the cells in the microwells. A representation of the glass microarray is shown in Figure 3 (a,b). The correlation between the number of cells secreting per well and fluorescence intensity can be done by imaging the PDMS slab to identify the number of cells per well (Figure 3, c).

There are three main advantages of microengraving over other traditional screening techniques such as serial dilution and ELISA. First, microengraving allows for the identification and segregation of the cells that secret antigen-specific antibodies from a polyclonal mixture early in the screening process, as opposed to serial dilution, for example, that requires outgrowth of the fast-growing subpopulations. Second, cells are segregated early in the process so only the desired ones will follow on further characterization. This reduces the labor and time required to maintain many individual

clones while characterizing the antibodies produced for appropriate reactivity in immunochemical assays. Third, microengraving simplifies the requirements for screening polyclonal populations to identify clones with different specificities. Multiple differentially labeled antigens can be simultaneously screened in a single microarray, whereas in ELISA, for example, an equivalent analysis would require independent assays for each condition[62]. Compared to flow cytometry, analysis of protein by the latter requires a physical link between phenotype and genotype either by surface capture[48] or display[66].

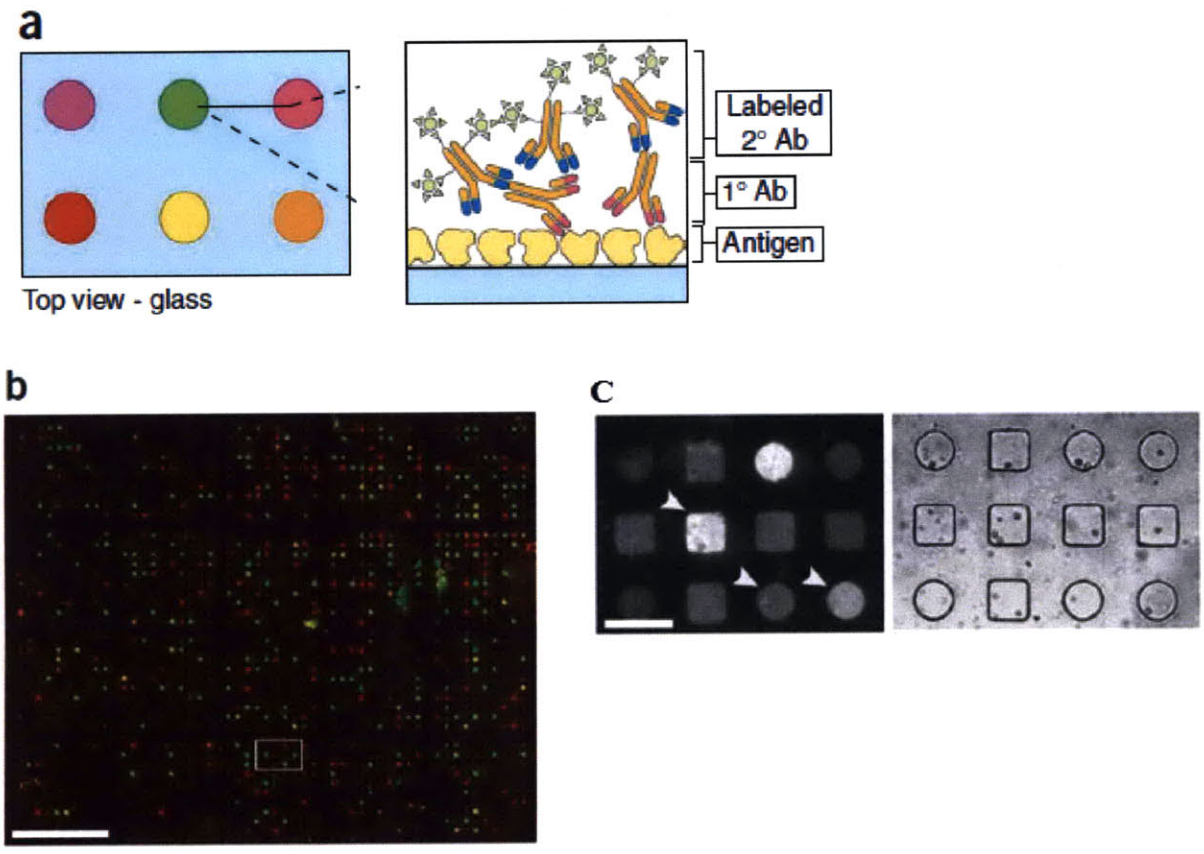


Figure 3: a) Antigen-coated slides capture secreted antibody that is detected with labeled secondary antibody. b) Micrograph of a microarray prepared using microwells (50-mm diameter). Scale bar, 1 mm [1]. c) fluorescence migrograph of an array. The arrowheads indicate spots where the differences in measured intensities arise from the different numbers of cells per well (three compared to one) or different amounts of antibody produced by single cells (two round wells). Scale bar, 200 μm

Arguably, the two strongest advantages for using microengraving in clonal selection is first, the ability to retrieve desired single cells, with an automated robotic tool (Cellcelector, Aviso GmbH), and second the ability to make measurements over time on the same set of cells, which gives a more accurate representation of dynamic processes, like protein secretion, that occur in cells.

## 2.3 *Pichia Pastoris* as a host organism for heterologous protein secretion

Following the discussion in Chapter 1 about different cell types used as hosts, *P. pastoris* is a good choice of secretion host for two main reasons. First, the production of recombinant proteins with this organism is relatively cheap. Second, this yeast type has been engineered to perform all the necessary post-translational modifications required for a protein to be functional[3, 21]. Thus, *P. pastoris* is the host organism we used for this project. The particular strain that was used to quantify protein production secretes a human Fc antibody fragment under the control of the strong, constitutive GAPDH promoter[67]. This promoter was chosen in an attempt to focus on the intrinsic secretory diversity among a population of cells, rather than promoter-induced differences. Additionally, the GAPDH promoter is commonly used for the expression of many proteins[68].

### 2.3.1 Dynamics in P. pastoris protein secretion over time

The first approach to characterize the *P. pastoris* population was done by tracking the secretion of protein over time and relating productivity to cell growth. Cells were grown for 4 days in YPD (Yeast Peptone Dextrose) media and samples were taken from the culture at different time points. At each time point, the optical density (OD) of the culture was

measured at 605 nm wavelength to determine the number of cells present. As seen in Figure 4, the cells followed an expected exponential growth curve, where they reach a stationary phase after a few days of divisions. To measure the amount of antibody secreted in the culture as a function of time, samples of media were taken at different time points and were analyzed by ELISA (Enzyme-linked Immunosorbent assay). The OD-corrected values of antibody secretion, as well as the titer of the culture at different time points, are shown in Figure 5. The secretion in this particular strain reaches a peak in production after which the culture productivity slows and eventually even declines, likely as a function of cell lysis and release of proteolytic enzymes capable of degrading the Fc present in the culture media.

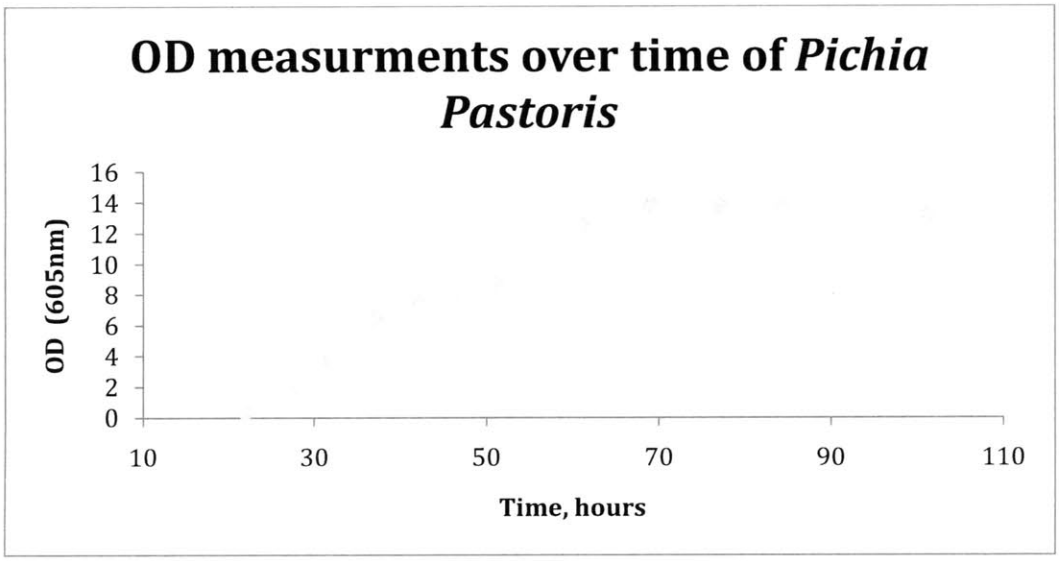


Figure 4 : OD measurement of *Pichia Pastoris* culture, under GAPDH promoter, over 5 days

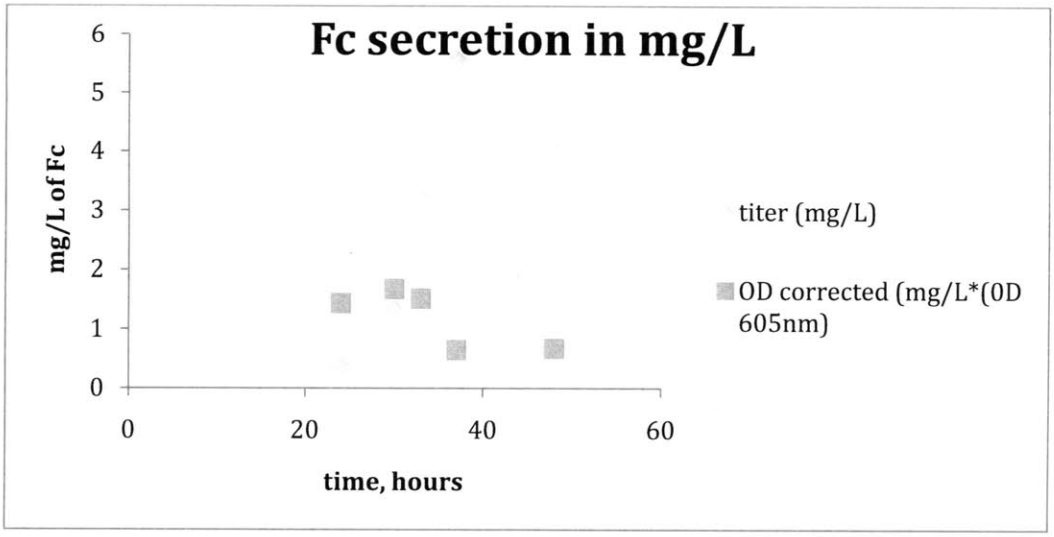


Figure 5: Secretion of Fc over time. The OD corrected values and the titer data are shown for P. Pastoris. Population

## 2.4 "Fishing" for high producers through clonal selection

In searching for high producing cells, it is useful to understand the diversity of subpopulations with respect to secretion in the culture. Studying the heterogeneity present can provide information about the secretion behavior of the culture, in terms of how much product is actually secreted by the population and whether the secretion is uniform for all cells.

The secretion rates in a wild type *P. pastoris* population secreting a human Fc fragment under the control of the strong, constitutive promoter GAPDH were measured using microengraving. The experiment was conducted for 1hr as described in Figure 7,a. The rate of secretion was determined by the fluorescence intensity presented on the glass slide, and a standard curve as shown in Figure 8. The number of cells per well was identified by imaging the PDMS array using a high-speed inverted fluorescence microscope. The distribution of secretion rates from single cells, as well as the percentage of secreting or

non secreting subpopulations within the entire population, is shown in Figure 9. Much heterogeneity in production exists in this population. The very top producing cells (marked with a yellow circle in Figure 9) from a parent population of Fc-secreting *P. pastoris* were retrieved and further characterized to determine if the phenotypes for high secretion remain stable after multiple rounds of division. If this was true, one would expect the new clonal populations to have an overall improved productivity and the distribution curve for single-cell rates of secretion to be shifted towards higher values. Again, using the microengraving technique the distribution of the secretion rate of these new clonal populations was determined (Figure 10a). The population originating from a single high producing cell is identical to the parental population, indicating that high productivity is an epigenetic phenomenon that does not last for many generations[69].

To extend the above result, non-producing and medium producing single cells from the parent population (red and green circles in Figure 9) were also retrieved and regrown as clonal populations. The distributions in single-cell rates of secretion for these new populations do not dramatically change compared to the parent clone (Figure 10a).

To prove the validity of the distribution data shown above, supernatants from all high, medium low, and non-producing cultures were assayed with ELISA to determine the actual amount of Fc production. As seen in Figure 10b, all daughter cultures produced from the retrieved cells secrete comparable amounts of Fc.

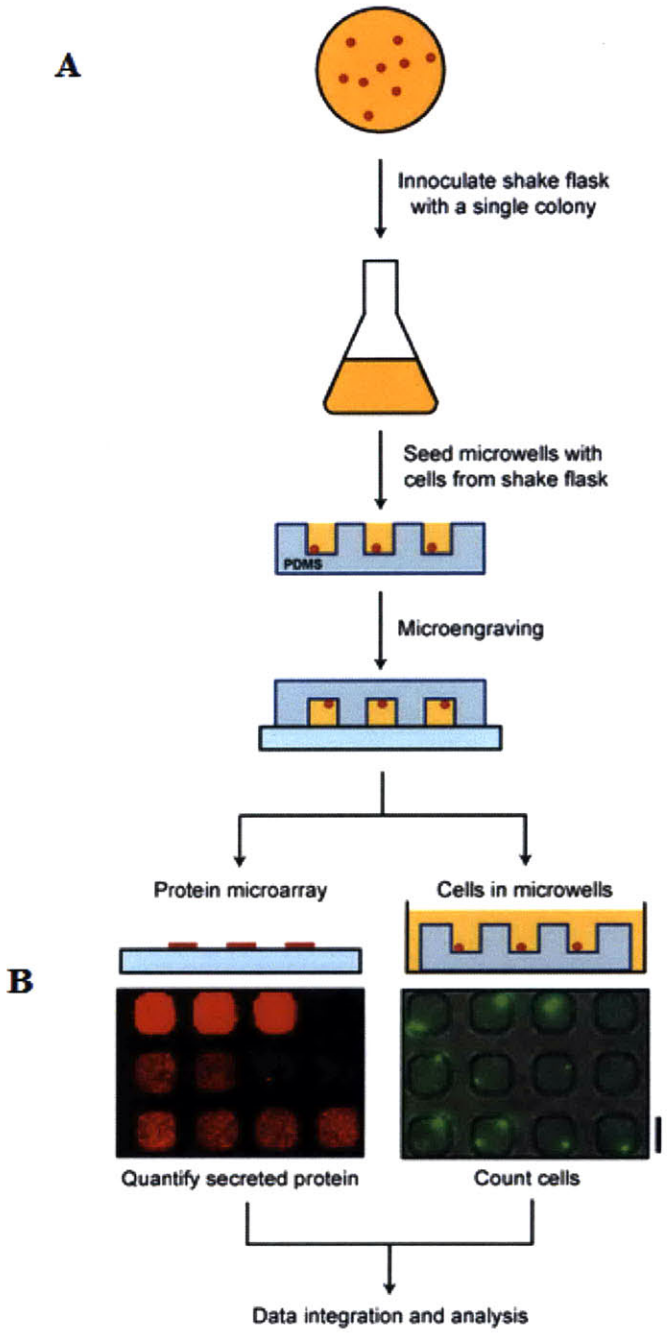


Figure 7: a)Microengraving with *P. pastoris* correlates secretion of heterologous proteins with single cells. A single colony of yeast cells was used to inoculate a shake flask, and grown for 1 to 2 days at 25°C. Cells from the shake flask were deposited onto an array of microwells at a density of ~1 cell/well. Microengraving was then conducted to create a protein microarray comprising the secreted proteins captured from each individual well. b)Secreted proteins were quantified and correlated with the cells present in the corresponding microwells. Scale bar is 50 μm[69]

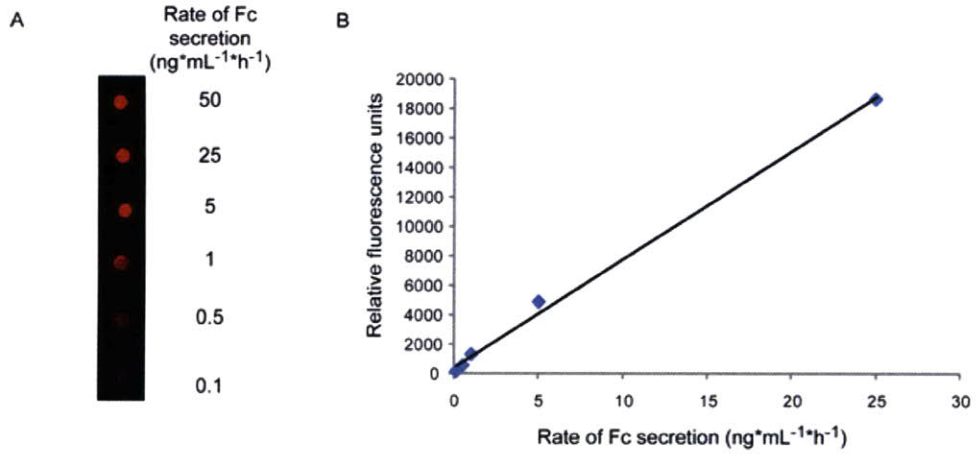


Figure 8 : Spot assay for conversion of the median fluorescence intensity (MFI) measured for captured Fc secreted by *P. pastoris* into a quantity of protein. **A:** Known concentrations of Fc were incubated in duplicate for 1h on a glass slide previously treated with goat anti-human Ig(H+L) (Zymax, Invitrogen). The slide was developed with Cy5-labeled goat anti-human Fc (Jackson) and scanned for fluorescence at 635 nm using a Genepix array scanner (Molecular Devices). **B:** Background-corrected fluorescence intensities for each concentration of Fc were used to construct a standard curve for the rate of Fc secretion ($ng*mL^{-1}*h^{-1}$). The data were fitted by linear regression ($R^2 = 0.996$).[69]

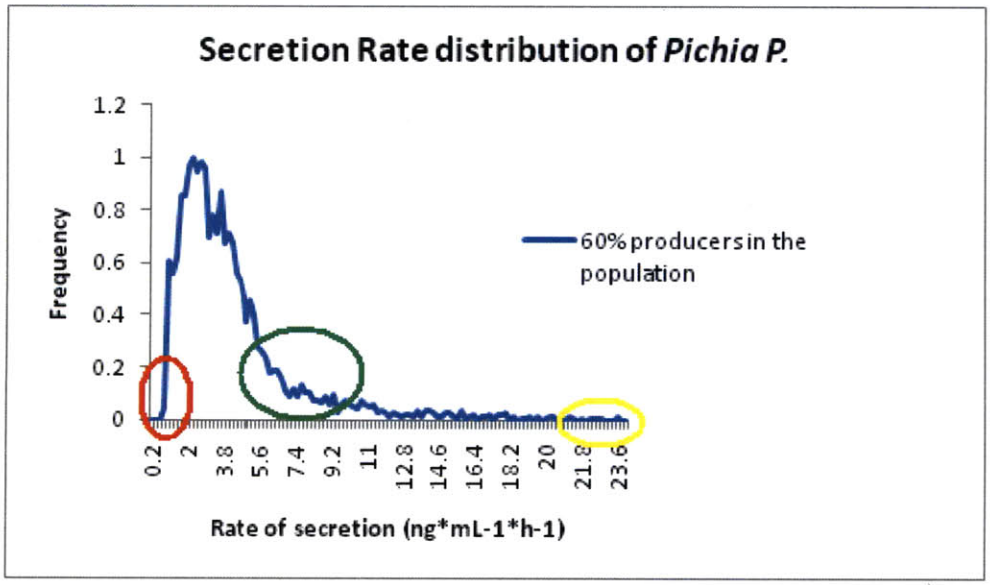


Figure 9: *Pichia pastoris* secretion rate distribution. The red circles shows the low producing area where cells were picked, the green shows the medium producers and the yellow the high

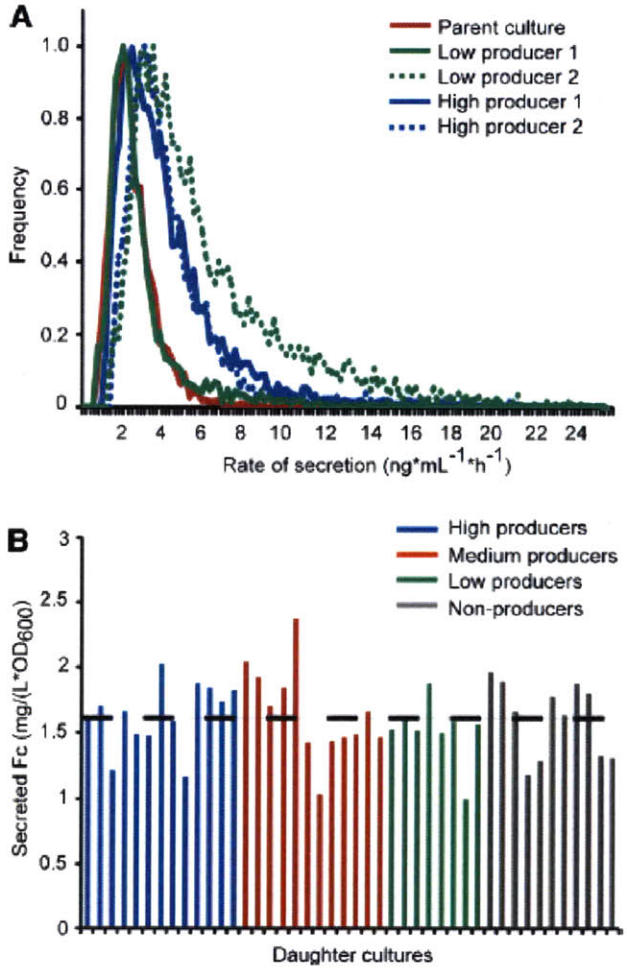


Figure 10 : Daughter culture secrete similar titers of Fc fragment at similar rates regardless of secretion status of parent cells. **A:** Histogram of the distributions of rates of secretion for single cells in clonal populations selected from a parent culture. Single cells from a clonal parent culture (red line) were identified as either high or low producers and were recovered and grown to late log phase in liquid culture. Single-cell distributions of rates of secretion for these new clonal lines were investigated via microengraving and are presented in duplicate (high producers, blue lines; low producers, green lines). Frequencies of secreting cells were the same for all clones shown (70±5%). **B:** Plot of secreted Fc fragment titer from liquid cultures following clonal expansion of single cells retrieved from a parent culture. Individual cells were selected based on secretion phenotype in the parent culture (blue, high producers; red, intermediate producers; green, low producers; gray, non-producers), grown to late log phase and supernatants were harvested. Bars represent an average of three replicate measurements of secreted Fc as determined by ELISA. Dashed line indicates the median titer for all 45 samples.[69]

## 2.5 Engineering a high producer

Since high productivity of a wild type clone is only an epigenetic effect that does not last for many generations, there is a need for alternative engineering strategies to make better producers[1]. We set out to increase secretion productivity by changing the internal environment of the cells with random mutagenesis. For this, the wild type strain of *P. Pastoris* was subjected to 3 rounds of chemical mutagenesis, using EMS (Ethyl Methyl Sulfonate) as the mutagen. Random mutagenesis has not been previously used for *P. pastoris* strain improvement due to the challenges associated with screening for desired phenotypes. Also, strain improvements gained via random mutagenesis often cannot be transferred to other strains[70]. We believed that microengraving would provide a unique method for simultaneously screening the entire mutagenized population based on the phenotype of secretion; thus, it was used in this work as our screening technique.

Directed evolution is a widely-used strategy to improve the stabilities or biochemical functions of proteins by repeated rounds of mutations[71]. Although directed evolution experiments may vary in details, they all use the same paradigm illustrated in Figure 11. Each experiment starts with a parent clone or a gene of interest that is mutagenized to produce a library of mutant progeny. The second step in the process is the screening of progeny phenotypes. The clones with desirable behavior are selected based on pre-deterimed criteria and used as parent clones for the next round. The success of the experiment depends on the feasibility of the target and whether measurable improvements can be accumulated in each round. After each round of mutagenesis the percentage of beneficial, neutral, or deleterious mutations is typically the same; 0.5-0.01%, 50-70% and

30-50% correspondingly[71]. The very small number of beneficial mutations indicates that the screening technique must be extremely sensitive.

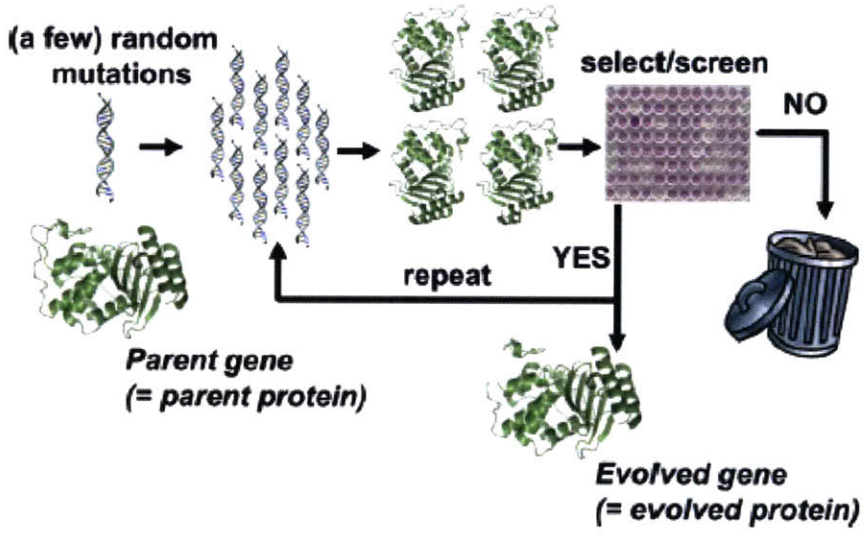


Figure 11: Schematic outline of a typical directed evolution experiment. A gene(s) for the parent protein(s), is randomly mutagenized and the library of mutant genes is then used to produce mutant proteins, which are screened or selected for the desired target property Mutants that fail to show improvements in the screening/selection are typically discarded, while the genes for the improved mutants are used as the parents for the next round of mutagenesis and screening. This procedure is repeated until the evolved protein(s) exhibits the desired level of the target property (or until the student performing the experiments graduates).[71]

In the case of *P. pastoris,* the mutation algorithm followed was the same as the one described above. The cells were mutated with EMS and subsequently analyzed by microengraving to identify the clones that would be used for the next round of mutagenesis. Based on published data, mutagenesis typically is successful when 60% of the population is killed[72]. To find the exact amount of EMS needed for a 60% kill rate, a killing assay was performed as shown in Figure 12.

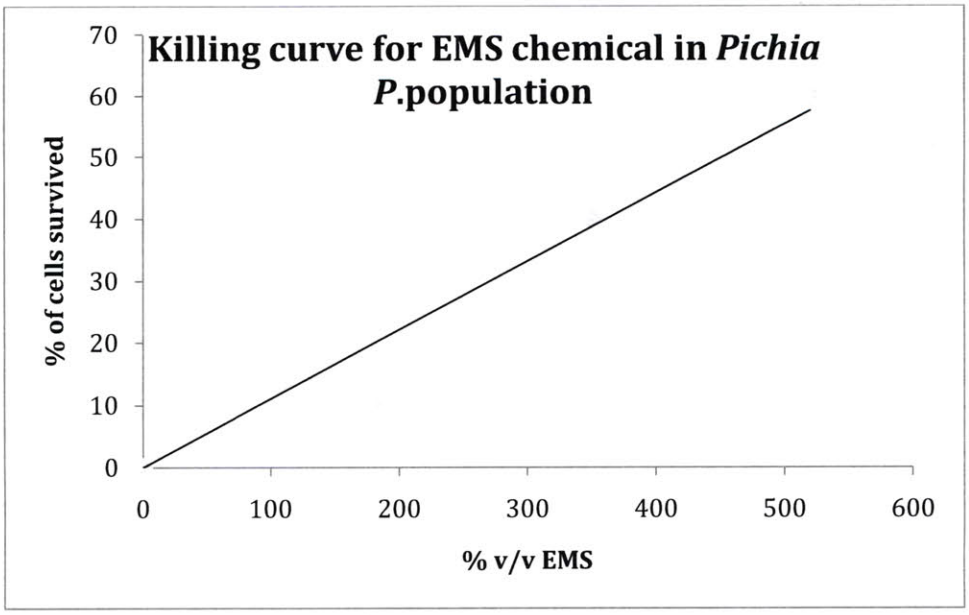


Figure 12 : Dose-dependent killing for EMS-treated *P. pastoris.* . Varied dosages of EMS were applied to a wild-type P. Pastoris. population and the cells were plated on YPD plates. The number of cells that survived each dose were counted by hand following 2-3 days of outgrownth and were compared to a control population of *P. Pastoris* where no mutagen was applied.

Three rounds of chemical mutagenesis were performed using enough EMS to produce a 60% kill in each round, and at the end of each the top secreting cells were retrieved. The rate of secretion for each single cell of the mutated population was determined by microengraving as described above, and data from all three mutagenesis rounds are shown in Figure 13,A. Based on the literature reports, we should expect that only 0.5-0.01% of the mutated population has beneficial mutations. This number corresponds to roughly 20-30 cells out of the more than 4000 single cell events that can be measured by microengraving. In order to follow strict criteria of selection, only the top 10 clones were retrieved and then carried forward as the parental population for the next round. To verify that each mutation round did indeed create diversity that improved secretion, all clones picked from that round were pooled and the distributions of secretion

rates for this polyclonal population were determined (Figure 13,B). Each polyclonal mutated population then served the basis for next round of mutagenesis. The goal of the experiment was to achieve a better population compared to the parental wild type population.

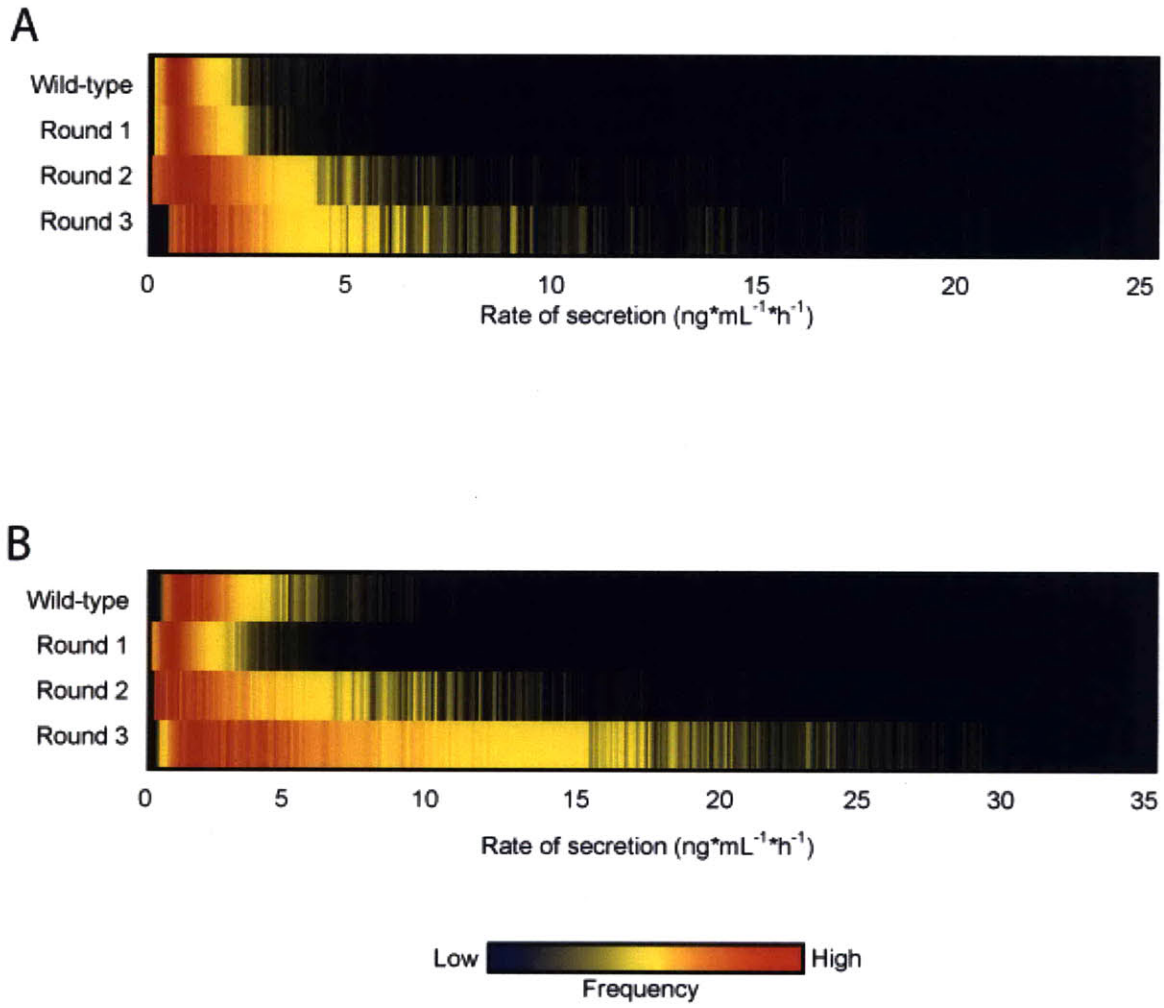


Figure 13 : Heat map representation of distributions of single-cell rates of secretion for *P. pastoris* A) Secretion rates directly following each of three rounds of chemical mutagenesis. Cells were grown to stationary phase and mutagenized with EMS. After 24 hrs of recovery the cells were loaded on the PDMS array and the microengraving experiment was conducted for 1 hr. At the end of the experiment the secretion rates of the mutagenized cells were measured. The control is the parental *P. pastoris* population prepared for mutagenesis with all the water dilutions and same recovery time, but without adding the EMS. B)Secretion rates after chemical mutagenesis experiments. The clones that were retrieved after each round were grown from single cells for 48hrs in 96-well plates and then used to innoculate 10ml cultures. Cultures were then grown to late log phase (OD~3-4), mixed together and the distribution rate of the polyclonal population was measured by microengraving. All 3 rounds are shown in this heat map. The control is the parental *P. pastoris* population.

As seen in Figure 13,B, the first round of mutagenesis appeared to have more deleterious mutations than beneficial ones, as the median single-cell rate of secretion for the mutagenized population was lower than the parental population. Following three rounds of mutagenesis, however, the population shows an overall improvement and contains many more high producing cells compared to the original population.

To verify the validity of these data, ELISA assays were performed on each individual clone from the final (third) round of mutagenesis. The retrieved best producing clones were individually grown in 1, 10, 50 & 500 ml cultures sequentially and supernatant samples from all culture volumes were collected at points both in early (OD ~ 4) and late (OD ~ 12) times of culture growth. The amount of human Fc secreted for each clone was determined and is presented in Figure 14. Values for each clone were normalized using the values obtained from the original parent *P. pastoris* population grown under the same conditions. As seen, most of the clones selected in the final round of mutagenesis are better secretors compared to the original culture early on in cultivation, but less productive as the culture reaches stationary phase. This probably indicates that the experimental growth conditions that were used for the microengraving in each round of mutagenesis (i.e., 24 hour outgrowth post-mutagenesis to ~OD 3-5 prior to microengraving) do affect the phenotype of the clones retrieved in each round.

We also noticed that when cells were grown up in the 500 ml media, they were less productive than when grown in smaller culture volumes (Figure. 14). This, together with the lack of productivity at higher culture densities, could indicate either overproduction of proteins inhibiting the further secretion of additional Fc or simultaneous production of other degrading or toxic proteins that affect the quantity of Fc already existing in the

culture supernatant. To test the role of other secreted factors on culture titer, a "conditioned media" experiment was conducted. Cells from each of the three best-producing clones were grown from frozen stocks, along with the wild-type strain, in 50ml of YPD to stationary phase. A 10 μL aliquot of each was then used to inoculate a new culture, where the outgrowth media was the supernatant of the outgrown culture treated with an additional supply of carbon source. Another small portion of the same cells were grown in fresh media and were used as a control. Supernatants from all cultures were sampled both early and late in culture growth and were assayed by ELISA for Fc titer (Figure 15).

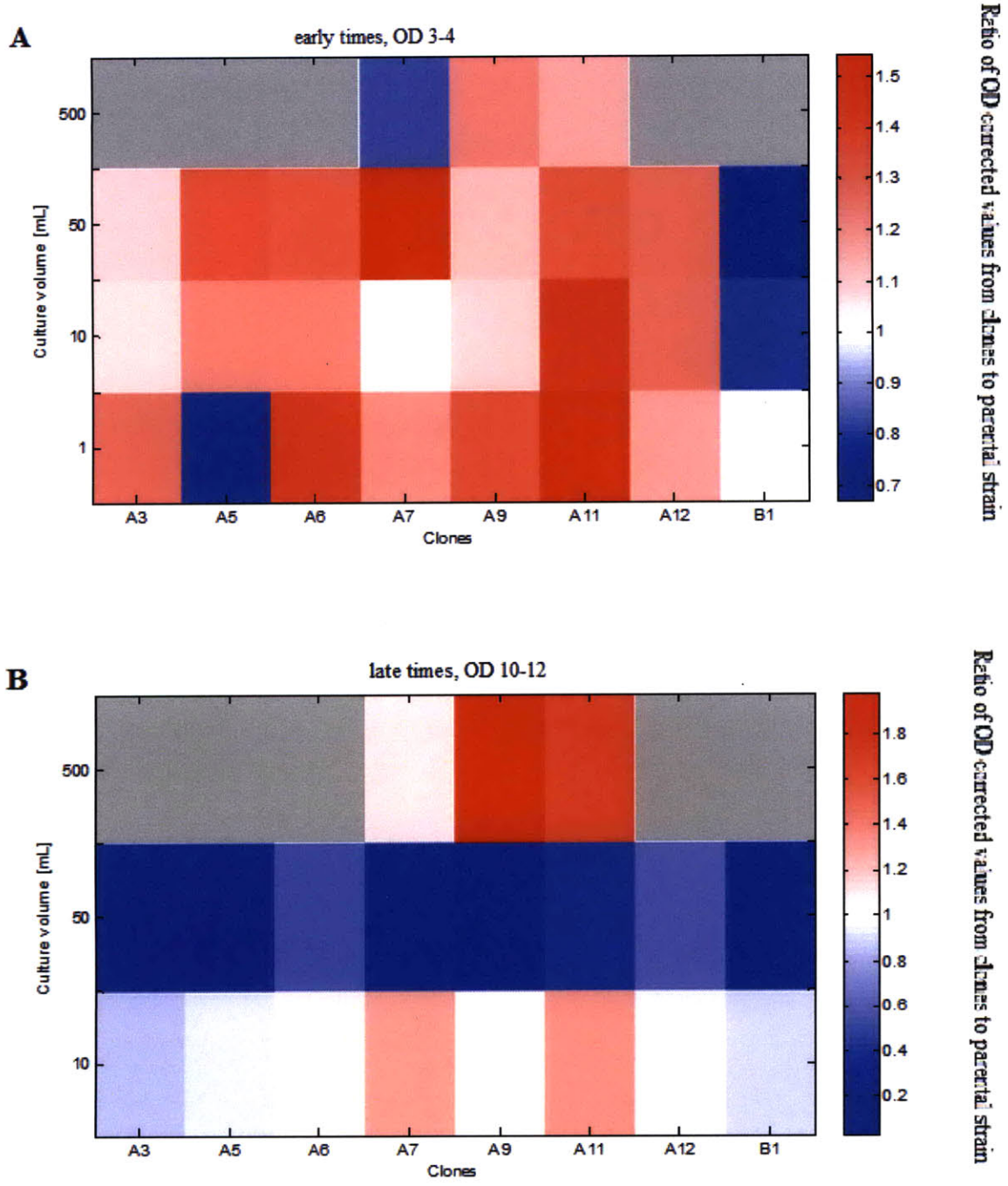


Figure 14: Fc titers as determined by ELISA on top producing clones from the 3rd round of mutagenesis grown in various culture volumes. The OD corrected values are shown for samples collected (A) early in culture growth (OD~3-4)and (B) late in culture growth(OD~10-12). No measurements were done on grey squares.

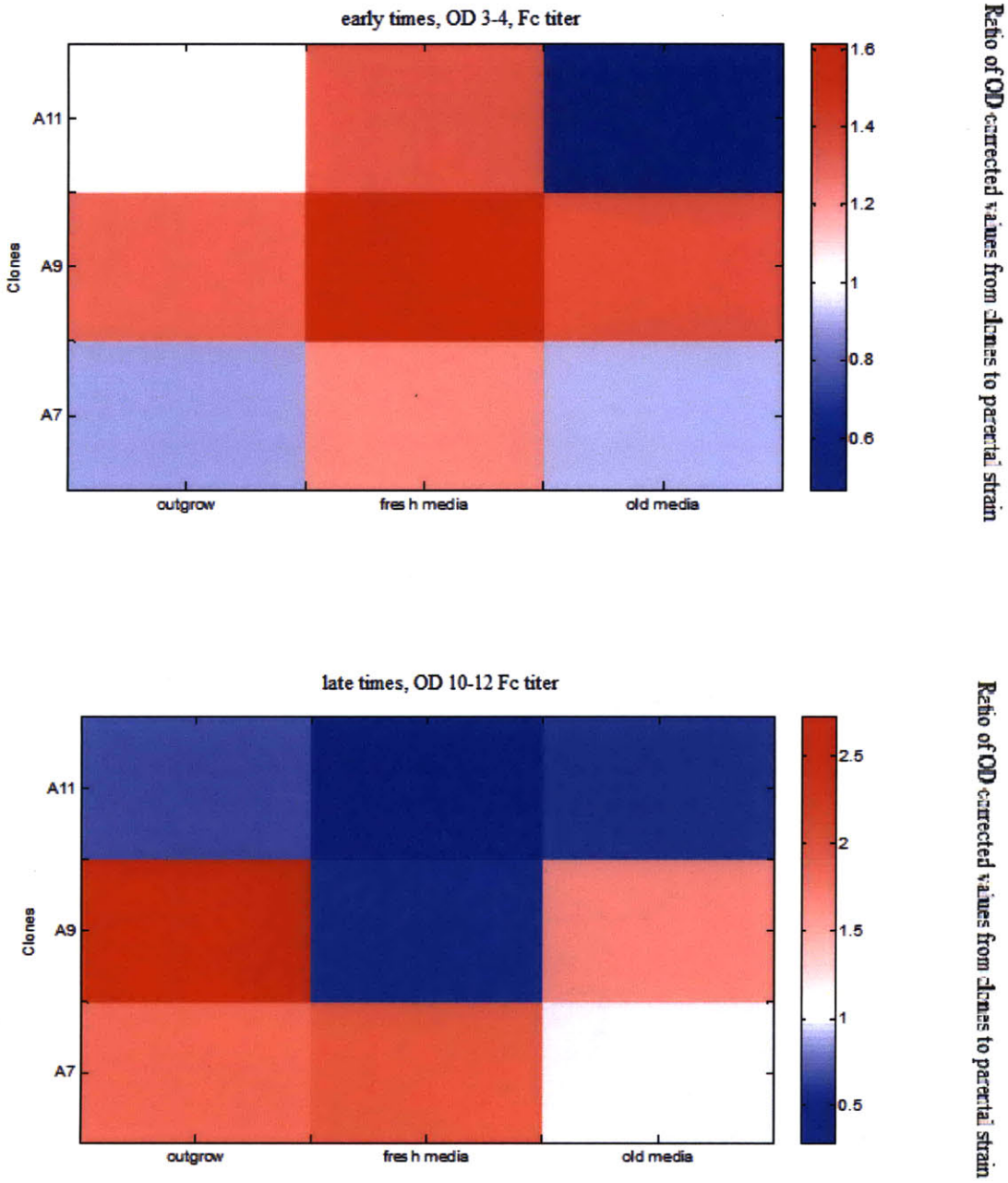


Figure 15: Overall Fc titers normalized using wild-type strain Fc titer prior to and following outgrowth in conditioned media. The three top producing clones were used to inoculate 50 ml of YPD and the supernatants were sampled in both early and late times in culture growth(first column on the left). The cells were harvested following an initial growth phase and the supernatant was retained. 10µl of cells from each culture were used to inoculate the corresponding supernatant. Additional glucose was added to each culture as a carbon source. Supernatants again were sampled both early and late in culture growth (last column on the right). Another 10µl of each culture was inoculated into fresh YPD grown to stationary phase. Supernatants were sampled both early and late in culture growth for comparison(middle column).

## 2.6 Discussion and Conclusions

In this chapter, we have shown a way to improve the secretion of an Fc fragment human antibody in *P. pastoris* population. First, we proved that high productivity within an isogenic population is an epigenetic effect that does not last for many generations. In order to improve the properties of the culture, engineering techniques that affect the genome need to be applied. In this project, random chemical mutagenesis was applied to a parent culture and microengraving was used to screen for high producers. Three rounds of chemical mutagenesis were applied to the culture, and at the end of the last round, the mutated population had 60% increase in the secretion rate. The validity of these data was proved by ELISA on each of the single clones from the last round of mutagenesis. As seen in Figure 15, almost all cultures from single clones retrieved from last round have higher titers than the parental clone. This though, is only true when the culture is sampled early in growth (when the OD of the cultures is relatively low). Later on, data appear inconsistent and clones do not seem to be better producers than the parental clone throughout the whole volume expansion. These results indicate two things for consideration. First, it could be that as the mutant clones grow they produce other proteins as well, which are potentially toxic. This might prevent the cells from further growth and degrade any existing proteins in the media. Supernatant sampling following culture growth in conditioned media shows that this could be happening: cells do not secrete additional Fc in to the culture media even though additional nutrients are supplied. Another likely factor is the fact that we screen for high producers at the point in growth of an OD around 3-4. As seen many times over in the literature, we are indeed identifying that for which we screen: cells retrieved are better producers on at the same ODs used in screening, but not as good as the culture continues to grown since we did not screen cells at this point in growth.

## 2.7 Materials and methods

***P. pastoris* cell culture.** *P. pastoris* secreting a human Fc under the GAPDH promoter (a gift from GlycoFi, Inc.) was streaked from a frozen clonal stock onto solid YPD (Yeast Peptone Dextrose) media. Colonies were allowed to develop at 25°C for several days. A single colony was used to inoculate 10 mL liquid YPD and the culture was then grown to late log phase ($OD_{600}$ = ~1.0-1.5) at 25°C with shaking at 290 rpm.

**Microengraving using *P. pastoris*.** Microwell arrays containing ~85,000 wells (each 50 μm in diameter) were fabricated in poly(dimethyl-siloxane) (PDMS) as reported previously using photolithography and replica molding [73]. PDMS arrays were sterilized by treatment with an oxygen plasma (PDC-32G, Harrick) for 20s and then were submerged in a 1% aq. solution of (3-glycidoxypropyl)trimethoxysilane for 50 min at 80°C. PDMS arrays were washed twice with sterile PBS before depositing cells. *P. pastoris* cells were harvested from liquid culture by centrifugation, resuspended in PBS at a density of 1 X $10^6$ cells/mL and stained with Calcofluor dye (250 μg/mL) for 10 min at 25°C. Cells were deposited onto the treated microwells arrays by dispensing 500 μL of the cell suspension dropwise onto the array. The cells were allowed to settle into the microwells for 5 min before excess media was removed and cell deposition was repeated. The surface of the array was dewetted by aspiration and washed with YPD. (Cells cultured in these wells were viable and expanded when left overnight at 25°C.) Glass slides were prepared as described [73] using 25 μg/mL goat anti-human Ig(H+L) antibody (Zymed) as the primary antibody for Fc capture. The array of microwells filled with *P. pastoris* cells was dewetted of excess media and was placed well-side down onto the surface of a treated, dry glass slide. The sandwich comprising the array and the glass slide was held together in a hybridization chamber (G2534A, Agilent Technologies). The entire assembly was incubated at 25°C for 1h. After incubation, the treated glass slide was removed from the surface of the microwell array and was immediately immersed in PBS.

**Interrogation of the printed microarray.** Following microengraving, glass slides were blocked (1% milk/0.05% PBS Tween 20) for 30 min and washed as described [73]. The slide was then treated with a solution of goat anti-human IgG secondary antibody (Cy5 conjugate, Jackson ImmunoResearch) at 0.5 ng/mL and incubated for in the dark for 45 min at 25°C. Slides were washed, dried, and then imaged with a microarray scanner (GenePix 4200AL, Molecular Devices) using a 635-nm laser and factory installed emission filters. The laser was used at 70% power and the PMT gain was set at maximum 450; these values were determined to maximize the linear range of detection in these experiments.

**Spot assay for construction of standard curve.** The fluorescence intensity for each individual spot on the engraved protein microarray was converted to a quantity of protein using a standard curve. The standard curve was obtained by constructing a protein array using known quantities of Fc (50, 25, 5, 1, 0.5 and 0.1 ng/mL) diluted in YPD and spotted on a glass slide as treated above. The slide was incubated for 1h, then developed and imaged as described above. Background-corrected fluorescence values were plotted against concentration to determine the linear range of the microengraving assay.

**Microscopy and micromanipulation.** Phase contrast and fluorescence images of the cell-loaded PDMS microarray were acquired using AxioVision software (v4.7.2, Carl Zeiss) and an automated inverted microscope (AxioObserver Z1, Carl Zeiss) equipped with a Hamamatsu EM-CCD camera. *P. pastoris* cells were retrieved from individual wells using a CellCelector (Aviso) as described [74] with optimized settings for yeast cells. Retrieved cells were each deposited into a separate well of a 96-well plate containing 200 µL YPD. 96-well plates were incubated at 25°C for 1-2 days before using the contents of a well to inoculate a larger liquid culture (1-10 mL YPD).

**Data processing and statistical analysis.** Phase contrast and fluorescence images of the cell-loaded PDMS microarray were analyzed for identification of the number of cells in each well using

the MabAnalyze program (custom script). Images of the printed microarrays were analyzed using GenePix Pro 6.0 (Molecular Devices). The background intensity for each array was determined from the median of all values measured in regions between individual spots of the array. Spots in the array were identified as positive when the signal-to-noise ratio was greater than 2--that is, when the spot intensity was greater than the sum of the background intensity for the array plus two standard deviations of the values used to calculate the background intensity. Multidimensional data were correlated using MatchBox (custom script). All subsequent data filtering and analysis was performed using Microsoft Excel.

**Random chemical mutagenesis of *P. pastoris*.** *P. pastoris* cells were grown to stationary phase in 10 mL YPD from a frozen stock. The population was diluted in water (1:10) and 300 µl of the EMS was added to the solution. The cells were incubated with shaking for 30 min and the mutagenesis reaction was stopped by addition of 10ml of 5% thiosulfate solution, followed by 5 min additional incubation. Cells were washed twice with YPD then allowed to recover in 10ml YPD for 24 hrs. Control cells were treated similarly with water dilution and shaking, but without the addition of EMS.

**Killing curve for determining optimal mutagen kill rate.** Different dosage of the EMS were applied to the same culture. Cells were treated as described above for chemical mutagenesis using concentrations of EMS at 0, 150, 350, 520 %v/v . After recovery, 10 µl of each different condition were plated on YPD platesand colonies were counted after 2-3 days.

**Fc titer determination by ELISA.** A 96 well plate treated for ELISA (Immulon 4 HBX, Nunc) was incubated with 5 µg/ml of goat-anti-human IgG(H+L) (Zymax, Invitrogen) at room temperature overnight. The plate was then blocked with 0.25% BSA in PBS/Tween for 30 min at 37 $^{0}$C. After blocking, the plates were dried and the supernatants of each clone were added in triplicates and diluted 5-fold, 6 times. In each plate, duplicates of the Fc standard, beginning with a concentration

of 200 ng/mL , were also added using 4-fold dilutions. 6 wells were intentionally left blank for background subtraction on each plate. The plate containing all dilutions of antigen was incubated at 37 °C for 30 min.

After incubation, the plate was washed 3 times with water and blocked with blocking buffer at room temperature for 10 min. Then, the detection antibody, goat-anti human Fc-alkaline phosphatase conjugated (Jackson Immunoresearch), was added at 75ng/mL and the plate was incubated for another 30 min at 37 °C. After this step, the plate was washed and blocked as before. Finally, 100μL of *p*-Nitrophenyl Phosphate, Disodium Salt (pNPP) substrate were added to each well and the color was left to develop at room temperature. The absorbance of each plate was read at 405 nm. Using the concentrations of Fc standard in each plate, the concentration of Fc in the supernatants of each culture was determined.

***P. pastoris* growth and Fc titer in conditioned media.** Frozen stocks of the clones were used to inoculate 50 ml of YPD in baffled flasks. Supernatants of the culture were sampled both at early and late times in culture growth, OD ~ 3-4 and 10-12 correspondingly. The cells were harvested when they reached stationary phase and the supernatant was kept and used as a media to re-inoculate 10μl of the same cells. Extra 5ml glucose was added to each conditioned media culture. As a control for this experiment, 10 μl of the cells at stationary phase were also used to inoculate 50ml of fresh YPD. Supernatants were sampled in early and late growth times for each fresh media culture and the conditioned media culture. An ELISA, as described above, was performed to determine Fc titer in all supernatants collected.

# 3 Chapter 3:

Discussion - Conclusions

## 3.1 Thesis Summary

For many years now, recombinant proteins produced by different host organisms have been used as therapeutics. Among the different hosts used in industry and academia so far, the yeast *P. pastoris* is particularly promising since the production of recombinant proteins from this organism is easy and it performs all the necessary modifications needed for a protein to be functional.

In this work, a wild-type strain of *P. pastoris* was used as a host organism to produce a fragment of the Fc human antibody. The purpose of the work was to improve the heterologous protein secretion of the host strain by enhancing the productivity of single cells. This task is challenging because the heterogeneity of the secretors within the same population makes screening for single producers difficult. We used microengraving to screen simultaneously a large library of *P. pastoris* cells individually for secretion capability.

The first experiments were performed to characterize the population distribution of the wild-type strain in terms of secretion. As was expected, the population has a burst in production after which it slows down. Single-cell heterogeneity in secretion among the population was further characterized by microengraving. The distribution of the secretion rates is substantial, but consistent with reports of clonal variation observed in other organisms, like *E. coli*[75] and CHO cells[76]. When the top producing clones, based on secretion, were isolated from the population and were regrown for many generations, the distribution of secretion rates remained similar to the parental population. The same phenomenon was seen for the low and medium producers as well. This phenomenon indicates that a cell's secretion productivity is an epigenetic effect that does not last for many generations.

Considering the above, alternatives for improvements in production were needed. For this, we chose to perform random chemical mutagenesis on the *P. pastoris* wild-type Fc-secreting strain.

After 3 rounds of chemical mutagenesis using EMS, a 60% better in secretion rate population was achieved. To verify this, all the clones from the last round of mutagenesis were assayed individually with ELISA, and the amount of antibody secreted was compared to that secreted from the original population. Titer measurements were done at both early and late growth times, and we noticed that when mutants grew to high ODs, they lost much of their productivity. This decrease is maybe due to the presence of other proteins that inhibit cell production and degrade existing proteins. To examine this possibility cells were grown to stationary phase and their supernatant was kept and used as media to regrow the same cells. The cells could not grow in this media even after adding extra carbon source, indicating the potential presence of inhibitors, or possibly other nutrient requirements.

In any case, an overall improved population based on production was made demonstrating a random mutagenesis methodology for enhancement of protein secretion and an advanced screening technique to identify high producing clones within a population

## 3.2 Discussion - Conclusions

In this thesis, I have demonstrated a methodology to improve the product secretion of a *P. pastoris* population by examining protein secretion at the single-cell level. The two challenges that are usually met in such experiments are first, the difficulty in screening because of the heterogeneity of the population and second optimization of a selection criterion that "pushes" the population toward improvement. Microengraving proved to be a great technique to address both of these challenges, since more than 4000 single cells could be simultaneously screened solely on the basis of secretion productivity and remain viable for further culture expansion. This number is significantly high considering other screening techniques, where the number of clones screened is in the order of 100s. Based on our initial experiments done with the wild type strain of *P. pastoris*, high levels of

secretion productivity is an epigenetic effect that does not last for many generations. In order to improve the secretion properties of a population, genetic changes must be made, so we applied chemical mutagenesis to the population of yeast cells.

One interesting result was that while the mutated population after round 3 did show an overall improvement of secreting population, this attribute was true only when cultures titers were compared early in growth. This result suggests that we got what we screened for, since all the screening following mutagenesis was conducted when the culture density was low.

Another observation was that the optimal clones at screening density tend to lose their productivity as they grow, which was seen even initially when following the un-mutated wild-type strain over time. If we extend the results of un-mutated wild-type strain to the mutated population, then the decrease in productivity at later times should have been expected, according to the behavior of the wild-type population. Furthemore, based on the data collected from the "conditioned media" experiment, cells could not grow on the spent media even when extra glucose was added. This result probably indicates the presence of inhibition factors in the media that prevents cells from growing and explains somewhat the inability to produce more proteins at late times. Another possibility is that some part of the clonal improvement could be attributed to epigenetics, which can take many generations to overcome. In other words, the improvement in productivity might be an epigenetic effect that is lost after many generations and duplication times.

Overall we conclude that random chemical mutagenesis is a useful technique for genome improvement for secretory function, but it requires an extremely sensitive screening method. Microengraving proved to be a unique tool; it is the only technique currently available to screen large numbers of single cells for real time protein production and it permits retrieval the desired cells alive for further characterization.

## 3.3 Future plans

While the methodology presented in this thesis gives successful results in terms of enhancing productivity, there are still improvements to be considered.

First, a few more rounds of random mutagenesis could follow to verify if the productivity can improve further, or if there is a limit at which the mutagenesis does not have any beneficial effects on the secretion rate of the population any more. In addition to that, alternatives in mutagenesis types could be tried, although initially in this work an effort was done using UV mutagenesis did not yield any promising results.

As far as the existing improved population is concerned, a further characterization is needed. We aim to understand why these cells secrete more, compared to the parental clone, and what genomic changes led to this phenotype. Initially, we will perform a transcriptional analysis of several of the identified clones, along with the wild-type strain under various growth conditions. This type of analysis will give information about genes that are differentially expressed in the new clones. If there are any significant differences in the expression of any gene, then it might be also useful to sequence the genome of a particular clone and identify any point mutations in it.

Finally, our selection criteria that led us to develop clones that are improved in production were strict in terms of selection as suggested by the directed evolution methodology. Keeping in mind that one will achieve what is screened for, an interesting experiment could be to allow mutagenized cultures to outgrow longer than allowed here before selection. In this way, we might achieve new *P. pastoris* variants that outperform the wild-type strain over a longer period of culture growth.

# 4 References


1. Browne, S.M. & Al-Rubeai, M. Selection methods for high-producing mammalian cell lines. *Trends in Biotechnology* **25**, 425-432 (2007).
2. Kim, N.S., Byun, T.H. & Lee, G.M. Key determinants in the occurrence of clonal variation in humanized antibody expression of CHO cells during dihydrofolate reductase mediated gene amplification. *Biotechnology Progress* **17**, 69-75 (2001).
3. Macauley-Patrick, S., Fazenda, M.L., McNeil, B. & Harvey, L.M. Heterologous protein production using the Pichia pastoris expression system. *Yeast* **22**, 249-270 (2005).
4. [Anon] Human Insulin Receives Fda Approval. *Clinical Pediatrics* **22**, 367-367 (1983).
5. Ferrer-Miralles, N., Domingo-Espin, J., Corchero, J.L., Vazquez, E. & Villaverde, A. Microbial factories for recombinant pharmaceuticals. *Microbial Cell Factories* **8**, - (2009).
6. Chartrain, M. & Chu, L. Development and production of commercial therapeutic monoclonal antibodies in Mammalian cell expression systems: an overview of the current upstream technologies. *Curr Pharm Biotechnol* **9**, 447-467 (2008).
7. Torphy, T.J. Pharmaceutical biotechnology - Monoclonal antibodies: boundless potential, daunting challenges. *Current Opinion in Biotechnology* **13**, 589-591 (2002).
8. Reichert, J.M. & Valge-Archer, V.E. Outlook - Development trends for monoclonal antibody cancer therapeutics. *Nature Reviews Drug Discovery* **6**, 349-356 (2007).
9. Potgieter, T.I. et al. Production of monoclonal antibodies by glycoengineered Pichia pastoris. *Journal of Biotechnology* **139**, 318-325 (2009).
10. Farid, S.S. Process economics of industrial monoclonal antibody manufacture. *Journal of Chromatography B-Analytical Technologies in the Biomedical and Life Sciences* **848**, 8-18 (2007).
11. Andersen, D.C. & Reilly, D.E. Production technologies for monoclonal antibodies and their fragments. *Current Opinion in Biotechnology* **15**, 456-462 (2004).
12. Li, F. et al. A systematic approach for scale-down model development and characterization of commercial cell culture processes. *Biotechnology Progress* **22**, 696-703 (2006).
13. Morton, C.L. & Potter, P.M. Comparison of Escherichia coli, Saccharomyces cerevisiae, Pichia pastoris, Spodoptera frugiperda, and COS7 cells for recombinant gene expression - Application to a rabbit liver carboxylesterase. *Molecular Biotechnology* **16**, 193-202 (2000).
14. Jayapal, K.R., Wlaschin, K.F., Hu, W.S. & Yap, M.G.S. Recombinant protein therapeutics from CHO cells - 20 years and counting. *Chemical Engineering Progress* **103**, 40-47 (2007).
15. Gerngross, T.U. Advances in the production of human therapeutic proteins in yeasts and filamentous fungi (vol 22, pg 1409, 2004). *Nature Biotechnology* **22**, 1589-1589 (2004).
16. Gellissen, G. et al. Heterologous Protein-Production in Yeast. *Antonie Van Leeuwenhoek International Journal of General and Molecular Microbiology* **62**, 79-93 (1992).
17. Cézanne, P., Reff, T., Shoemaker, I.H. & Philadelphia Museum of Art. Paul Cézanne, two sketchbooks : the gift of Mr. and Mrs. Walter H. Annenberg to the Philadelphia Museum of Art. (Philadelphia Museum of Art, Philadelphia, Pa.; 1989).
18. Moir, P.W. Efficient beyond imagining : CIM and its applications for today's industry. (E. Horwood ;
Halsted Press, Chichester
New York; 1989).
19. Porro, D. & Branduardi, P. Yeast cell factory: fishing for the best one or engineering it? *Microbial Cell Factories* **8**, - (2009).
20. Li, P.Z. et al. Expression of recombinant proteins in Pichia pastoris. *Applied Biochemistry and Biotechnology* **142**, 105-124 (2007).

21. Wegner, G.H. Emerging applications of the methylotrophic yeasts. *FEMS Microbiol Rev* **7**, 279-283 (1990).
22. Macauley-Patrick, S., Fazenda, M.L., McNeil, B. & Harvey, L.M. Heterologous protein production using the Pichia pastoris expression system. *Yeast* **22**, 249-270 (2005).
23. Cereghino, J.L. & Cregg, J.M. Heterologous protein expression in the methylotrophic yeast Pichia pastoris. *FEMS Microbiol Rev* **24**, 45-66 (2000).
24. Cregg, J.M., Vedvick, T.S. & Raschke, W.C. Recent advances in the expression of foreign genes in Pichia pastoris. *Biotechnology (N Y)* **11**, 905-910 (1993).
25. Lin Cereghino, G.P., Sunga, A.J., Lin Cereghino, J. & Cregg, J.M. Expression of foreign genes in the yeast Pichia pastoris. *Genet Eng (N Y)* **23**, 157-169 (2001).
26. Menendez, J., Valdes, I. & Cabrera, N. The ICL1 gene of Pichia pastoris, transcriptional regulation and use of its promoter. *Yeast* **20**, 1097-1108 (2003).
27. De Schutter, K. et al. Genome sequence of the recombinant protein production host Pichia pastoris. *Nat Biotechnol* **27**, 561-566 (2009).
28. Mattanovich, D. et al. Genome, secretome and glucose transport highlight unique features of the protein production host Pichia pastoris. *Microb Cell Fact* **8**, 29 (2009).
29. Chen, D.C., Wang, B.D., Chou, P.Y. & Kuo, T.T. Asparagine as a nitrogen source for improving the secretion of mouse alpha-amylase in Saccharomyces cerevisiae protease A-deficient strains. *Yeast* **16**, 207-217 (2000).
30. Jenkins, N., Parekh, R.B. & James, D.C. Getting the glycosylation right: implications for the biotechnology industry. *Nat Biotechnol* **14**, 975-981 (1996).
31. Grinna, L.S. & Tschopp, J.F. Size Distribution and General Structural Features of N-Linked Oligosaccharides from the Methylotrophic Yeast, Pichia-Pastoris. *Yeast* **5**, 107-115 (1989).
32. Callewaert, N. et al. Use of HDEL-tagged Trichoderma reesei mannosyl oligosaccharide 1,2-alpha-D-mannosidase for N-glycan engineering in Pichia pastoris. *Febs Letters* **503**, 173-178 (2001).
33. Choi, B.K. et al. Use of combinatorial genetic libraries to humanize N-linked glycosylation in the yeast Pichia pastoris. *Proceedings of the National Academy of Sciences of the United States of America* **100**, 5022-5027 (2003).
34. Chen, D.Z., Wang, B.D., Chou, P.Y. & Kuo, T.T. Asparagine as a nitrogen source for improving the secretion of mouse alpha-amylase in Saccharomyces cerevisiae protease A-deficient strains. *Yeast* **16**, 207-217 (2000).
35. Sreekrishna, K. et al. Strategies for optimal synthesis and secretion of heterologous proteins in the methylotrophic yeast Pichia pastoris. *Gene* **190**, 55-62 (1997).
36. Shi, X.Z. et al. Optimal conditions for the expression of a single-chain antibody (scFv) gene in Pichia pastoris. *Protein Expression and Purification* **28**, 321-330 (2003).
37. Sarramegna, V., Demange, P., Milon, A. & Talmont, F. Optimizing functional versus total expression of the human mu-opioid receptor in Pichia pastoris. *Protein Expression and Purification* **24**, 212-220 (2002).
38. Cos, O., Ramon, R., Montesinos, J.L. & Valero, F. Operational strategies, monitoring and control of heterologous protein production in the methylotrophic yeast Pichia pastoris under different promoters: A review. *Microbial Cell Factories* **5**, - (2006).
39. Hartner, F.S. et al. Promoter library designed for fine-tuned gene expression in Pichia pastoris. *Nucleic Acids Research* **36**, - (2008).
40. Davis, R., Schooley, K., Rasmussen, B., Thomas, J. & Reddy, P. Effect of PDI overexpression on recombinant protein secretion in CHO cells. *Biotechnology Progress* **16**, 736-743 (2000).
41. Humphreys, D.P., Weir, N., Lawson, A., Mountain, A. & Lund, P.A. Co-expression of human protein disulphide isomerase (PDI) can increase the yield of an antibody Fab' fragment expressed in Escherichia coli. *Febs Letters* **380**, 194-197 (1996).

42. Zhang, W., Sinha, J., Smith, L.A., Inan, M. & Meagher, M.M. Maximization of production of secreted recombinant proteins in Pichia pastoris fed-batch fermentation. *Biotechnology Progress* **21**, 386-393 (2005).
43. Spencer, J. & Spencer, D. Mutagenesis in Yeast. *Yeast protocols: methods in cell and molecular biology*, 17 (1996).
44. Jansen, G., Hazendonk, E., Thijssen, K. & Plasterk, R. Reverse genetics by chemical mutagenesis in Caenorhabditis elegans. *Nature genetics* **17**, 119-121 (1997).
45. Lawrence, C.W. & Christensen, R. Uv Mutagenesis in Radiation Sensitive Strains of Yeast. *Biophysical Journal* **15**, A196-A196 (1975).
46. Hoogenboom, H.R. Selecting and screening recombinant antibody libraries. *Nature Biotechnology* **23**, 1105-1116 (2005).
47. Wan, L. et al. High-level expression of a functional humanized single-chain variable fragment antibody against CD25 in Pichia pastoris. *Applied Microbiology and Biotechnology* **81**, 33-41 (2008).
48. Mattanovich, D. & Borth, N. Applications of cell sorting in biotechnology. *Microbial Cell Factories* **5**, - (2006).
49. Kacmar, J., Zamamiri, A., Carlson, R., Abu-Absi, N.R. & Srienc, F. Single-cell variability in growing Saccharomyces cerevisiae cell populations measured with automated flow cytometry. *Journal of Biotechnology* **109**, 239-254 (2004).
50. Kromenaker, S.J. & Srienc, F. Stability of Producer Hybridoma Cell-Lines after Cell Sorting - a Case-Study. *Biotechnology Progress* **10**, 299-307 (1994).
51. Barnes, L.M., Moy, N. & Dickson, A.J. Phenotypic variation during cloning procedures: Analysis of the growth behavior of clonal cell lines. *Biotechnology and Bioengineering* **94**, 530-537 (2006).
52. Hohenblum, H., Borth, N. & Mattanovich, D. Assessing viability and cell-associated product of recombinant protein producing Pichia pastoris with flow cytometry. *Journal of Biotechnology* **102**, 281-290 (2003).
53. Herzenberg, L.A. et al. The history and future of the fluorescence activated cell sorter and flow cytometry: A view from Stanford. *Clinical Chemistry* **48**, 1819-1827 (2002).
54. Deere, D. et al. Flow cytometry and cell sorting for yeast viability assessment and cell selection. *Yeast* **14**, 147-160 (1998).
55. Williams, G.B., Weaver, J.C. & Demain, A.L. Rapid Microbial Detection and Enumeration Using Gel Microdroplets and Colorimetric or Fluorescence Indicator Systems. *Journal of Clinical Microbiology* **28**, 1002-1008 (1990).
56. Gray, F., Kenney, J.S. & Dunne, J.F. Secretion Capture and Report Web - Use of Affinity Derivatized Agarose Microdroplets for the Selection of Hybridoma Cells. *Journal of Immunological Methods* **182**, 155-163 (1995).
57. Weaver, J.C., McGrath, P. & Adams, S. Gel microdrop technology for rapid isolation of rare and high producer cells. *Nature Medicine* **3**, 583-585 (1997).
58. Diercks, A.H. et al. A microfluidic device for multiplexed protein detection in nano-liter volumes. *Analytical Biochemistry* **386**, 30-35 (2009).
59. Kartalov, E.P. et al. High-throughput multi-antigen microfluidic fluorescence immunoassays. *Biotechniques* **40**, 85-90 (2006).
60. Mann, C. Rapid isolation of antigen-specific clones from hybridoma fusions. *Nature Methods* **4** (2007).
61. Szaniszlo, P. et al. Scanning cytometry with a LEAP: Laser-enabled analysis and processing of live cells in situ. *Cytometry Part A* **69**, 641 (2006).
62. Love, J.C., Ronan, J.L., Grotenbreg, G.M., van der Veen, A.G. & Ploegh, H.L. A microengraving method for rapid selection of single cells producing antigen-specific antibodies. *Nature Biotechnology* **24**, 703-707 (2006).

63. Ostuni, E., Chen, C.S., Ingber, D.E. & Whitesides, G.M. Selective deposition of proteins and cells in arrays of microwells. *Langmuir* **17**, 2828-2834 (2001).
64. Kang, I.K., Ito, Y., Sisido, M. & Imanishi, Y. Gas-Permeability of the Film of Block and Graft-Copolymers of Polydimethylsiloxane and Poly(Gamma-Benzyl L-Glutamate). *Biomaterials* **9**, 349-355 (1988).
65. Lee, J.N., Jiang, X., Ryan, D. & Whitesides, G.M. Compatibility of mammalian cells on surfaces of poly(dimethylsiloxane). *Langmuir* **20**, 11684-11691 (2004).
66. Wentz, A.E. & Shusta, E.V. Novel high-throughput screen reveals yeast genes that increase secretion of heterologous proteins. *Applied and Environmental Microbiology* **73**, 1189-1198 (2007).
67. Waterham, H.R., Digan, M.E., Koutz, P.J., Lair, S.V. & Cregg, J.M. Isolation of the Pichia pastoris glyceraldehyde-3-phosphate dehydrogenase gene and regulation and use of its promoter. *Gene* **186**, 37-44 (1997).
68. Cregg, J.M., Cereghino, J.L., Shi, J.Y. & Higgins, D.R. Recombinant protein expression in Pichia pastoris. *Molecular Biotechnology* **16**, 23-52 (2000).
69. Love, K., Panagiotou, V., Jiang, B., Stadheim, T. & Love, J. Integrated single-cell analysis shows Pichia pastoris secretes protein stochastically. *Biotechnology and Bioengineering* **9999**.
70. Gasser, B., Sauer, M., Maurer, M., Stadlmayr, G. & Mattanovich, D. Transcriptomics-based identification of novel factors enhancing heterologous protein secretion in yeasts. *Applied and Environmental Microbiology* **73**, 6499 (2007).
71. Bloom, J. & Arnold, F. In the light of directed evolution: Pathways of adaptive protein evolution. *Proceedings of the National Academy of Sciences* **106**, 9995 (2009).
72. Lawrence, C., Nisson, P. & Christensen, R. UV and chemical mutagenesis in rev7 mutants of yeast. *Molecular and General Genetics MGG* **200**, 86-91 (1985).
73. Ogunniyi, A.O., Story, C.M., Papa, E., Guillen, E. & Love, J.C. Screening individual hybridomas by microengraving to discover monoclonal antibodies. *Nat. Protoc.* **4**, 767-782 (2009).
74. Choi, J.H. et al. Development and optimization of a process for automated recovery of single cells identified by microengraving. *Biotechnol. Prog. in press* (2009).
75. Cai, L., Friedman, N. & Xie, X.S. Stochastic protein expression in individual cells at the single molecule level. *Nature* **440**, 358-362 (2006).
76. Pilbrough, W., Munro, T.P. & Gray, P. Intraclonal Protein Expression Heterogeneity in Recombinant CHO Cells. *Plos One* **4**, - (2009).