**Maria Jardim Caramelo Dias Cabral**

Licenciada em Anatomia Patológica, Citológica e Tanatológica

# Phage Display as a Tool for Development of Novel Therapeutics for Breast Cancer

Dissertação para obtenção do Grau de Mestre em
Genética Molecular e Biomedicina

Orientador: Ana Barbas, PhD, iBET

Júri:

Presidente: Prof. Doutora Ilda Sanches, FCT-UNL
Arguente: Doutor João Gonçalves, FF-UL
Vogal: Doutora Ana Barbas, iBET

FACULDADE DE CIÊNCIAS E TECNOLOGIA
UNIVERSIDADE NOVA DE LISBOA

**Setembro 2014**

Phage Display as a Tool for Development of Novel Therapeutics for Breast Cancer

# Acknowledgements

Gostaria de agradecer à direção do iBET por me ter permitido realizar nas suas instalações o trabalho que deu origem a esta dissertação.

Agradeço à Doutora Ana Barbas, minha orientadora, por me ter acolhido como estudante de mestrado e me ter confiado um papel neste projeto; pela motivação, apoio, confiança e positivismo com que sempre me confrontou e por todo tempo que despendeu a transmitir-me conhecimento e a orientar-me.

Agradeço aos restantes membros do laboratório satélite da Bayer Healthcare Pharma – Joana Galvão, Inês Barbosa e Fernanda Spínola Rodrigues –, à Ana Raposo, à Maria Margarida Silva e ao Manuel Garrido, igualmente por todo o apoio e conhecimento que me transmitiram, por se interessarem no meu trabalho e ajudarem a vencer os obstáculos que surgiram no percurso.

Agradeço aos meus amigos que estiveram presentes e partilharam este ano comigo, que ajudaram a vencer a pressão, a esquecer e ultrapassar os momentos menos bons e a viver outros tantos melhores – ao Bruno, à Rita e em especial à Mariana por me aturar (quase) todos os dias.

Agradeço aos meus amigos que não puderam estar tão presentes, sobretudo pela minha ausência, mas que sei que me vão receber com a amizade, companheirismo e cumplicidade de sempre, como se nenhum tempo se tivesse passado desde o nosso último encontro.

Agradeço ao João por ser o meu outro lado, o meu mundo paralelo, para onde vou quando quero estar longe e descansar deste.

Agradeço, finalmente e mais que tudo, à minha família, pelo mimo e carinho que sempre guardam e por às vezes se esquecerem que já sou crescida, por me darem o mundo e me abrirem as primeiras portas a tudo o que tenho e o que fiz, e por me terem feito aquilo que sou.

**Obrigada!**

## Abstract

Phage display technology is a powerful platform for the generation of highly specific human monoclonal antibodies (Abs) with potential use in clinical applications. Moreover, this technique has also proven to be a reliable approach in identifying and validating new cancer-related targets. For scientific or medical applications, different types of Ab libraries can be constructed. The use of Fab Immune libraries allows the production of high quality and affinity antigen-specific Abs. In this work, two immune human phage display IgG Fab libraries were generated from the Ab repertoire of 16 breast cancer patients, in order to obtain a tool for the development of new therapeutic Abs for breast cancer, a condition that has great impact worldwide. The generated libraries are estimated to contain more than $10^8$ independent clones and a diversity over 90%.

Libraries validation was pursued by selection against BSA, a foreign and highly immunogenic protein, and HER2, a well established cancer target. Preliminary results suggested that phage pools with affinity for these antigens were selected and enriched. Individual clones were isolated, however, it was not possible to obtain enough data to further characterize them. Selection against the DLL1 protein was also performed, once it is a known ligand of the Notch pathway, whose deregulation is associated to breast cancer, making it an interesting target for the generation of function-blocking Abs. Selection resulted in the isolation of a clone with low affinity and Fab expression levels. The validation process was not completed and further effort will have to be put in this task in the future.

Although immune libraries concept implies limited applicability, the library reported here has a wide range of use possibilities, since it was not restrained to a single antigen but instead thought to be used against any breast cancer associated target, thus being a valuable tool.



**Keywords:**

Phage display; Breast cancer; Notch pathway; DLL1; Human immune Fab Library; Therapeutic antibodies

# Resumo

A tecnologia de phage display é uma poderosa plataforma para criação de anticorpos (Acs) monoclonais humanos altamente específicos com potenciais aplicações clínicas. Esta técnica tem provado ser uma abordagem confiável para identificação e validação de novo alvos relacionados com cancro. Diferentes tipos de bibliotecas de Acs podem ser construídas para aplicações científicas e médicas. As bibliotecas imunes de fragmentos Fab permitem produzir Acs específicos de elevada qualidade e afinidade. Neste trabalho, duas bibliotecas imunes de phage display de Fabs de IgG humanos foram geradas do repertório de Acs de 16 pacientes de cancro de mama, para a obtenção de uma ferramenta para desenvolvimento de novos Acs terapêuticos para esta patologia de grande impacto mundial. Estimou-se que as bibliotecas geradas contêm mais de $10^8$ clones independentes e uma diversidade acima de 90%.

Pretendeu-se validar as bibliotecas selecionando-as contra BSA, uma proteína altamente imunogénica, e HER2, um reconhecido alvo de cancro. Resultados preliminares sugerem que ocorreu seleção e enriquecimento de pools de fagos com afinidade para estes antigénios. Isolaram-se clones individuais, no entanto não foi possível obter informação suficiente para os caracterizar mais profundamente. Também se realizou seleção contra DLL1, dado ser um conhecido ligando da via Notch, cuja desregulação está associada ao cancro de mama, tornando-a um alvo de interesse para criação de Acs bloqueadores de função. O resultado foi o isolamento de um clone com baixa afinidade e baixos níveis de expressão de Fabs. O processo de validação não foi concluído, pelo que futuramente se deverão investir esforços nessa tarefa.

Embora o conceito de biblioteca imune implique uma aplicabilidade limitada, a biblioteca aqui descrita tem vastas possibilidades de uso, pois não está restringida a um único antigénio; pelo contrário, foi idealizada para ser usada contra qualquer alvo associado a cancro de mama, sendo portanto uma ferramenta valiosa.



**Palavras-chave:**

Phage display; Cancro de mama; Via Notch; DLL1; Biblioteca imune de Fabs humanos; Anticorpos terapêuticos

# Contents

# List of Figures

# List of Tables

## Abbreviations

2YT – 2x yeast extract and tryptone broth
λ – Lambda
κ – Kappa
Ab – Antibody
ADAM – A Disintegrin And Metalloproteinase
amp/glu – 100 µg/mL ampicillin and 1% glucose agar plates
bp – Base pairs
BSA – Bovine serum albumin
C-terminal – Carboxy terminal
CBF1 – C-promoter binding factor 1
CDR – Complementarity-determining region
cfu – Colony forming units
CH – Constant heavy
CL – Constant light
CSCs – Cancer stem cells
CSL – CBF1, Su(H) and Lag1
Da - Daltons
cDNA – Complementary DNA
dH2O – Distilled water
DLL1 – Delta-like 1
DLL3 – Delta-like 3
DLL4 – Delta-like 4
DNA - Deoxyribonucleic acid
dNTPs - Deoxynucleotide triphosphates
DPBS – Dulbeccos's phosphate buffered saline
DPBST-MP – DPBS with 0,05% Tween 20 and milk poder (usualy 3%)
DSL – Delta/Serrate/lag-2
EBNA6 - Epstein–Barr virus nuclear antigen 6
EDTA – Ethylenediamine tetraacetic acid
EGF – Epidermal growth factor
EGFR – Epidermal growth factor receptor
ELISA – Enzyme-linked immunosorbent assay
EMA – European Medicines Agency
ER – Estrogen receptor
Fab – Fragment antigen binding
Fc – Fragment crystallizable
FDA – Food and Drug Administration Agency
FR – Framework region
Fw – Forward
GSIs – Gamma (γ) secretase inhibitors
HA – Hemagglutinine
HC – Heavy chains
HEK293 – Human Embryonic Kidney 293 cells
HES1 – Hairy and enhancer of split-1
HES5 – Hairy and enhancer of split-5
HER2 – Human epidermal growth factor receptor 2
HER2+ – HER2 positive
HEY2 – Hairy/enhancer-of-split related with YRPW motif protein 2
$His_6$ – Hexa-histidine tail
IPTG – Isopropyl-beta-D-thiogalactopyranoside
kb – Kilo base pairs
kan/glu – 50 µg/mL kanamycin and 1% glucose agar plates
kDa – Kilo daltons
LB – Luria Bertani broth
LC – Light chain
Igs – Immunoglobulin
mAb – Monoclonal antibody
MOI – Multiplicity of infection
MQ – Milli-Q purified water
mRNA – Messenger RNA
mTOR – Mammalian target of rapamycin
N-terminal – Amino terminal
NICD – Notch intracellular domain
NRARP – NOTCH-regulated ankyrin repeat protein
NRR – Negative regulatory domain
O/N – Overnight
PBST-MP – Phosphate buffered saline with 0,05% Tween 20 and milk poder (usualy 3%)
PCR – Polimerase chain reaction
PR – Progesterone receptor

RNA – Ribonucleic Acid
rRNA – Ribossomal RNA
RT – Room temperature
Rv – Reverse
scFv – Single chain variable domains
sFab – Soluble Fab
TAE – Tris-Acetate-EDTA
T-ALL – T-cell lymphoblastic leukemias
TEA – Triethanolamine
TMB – 3,3',5,5'-tetramethylbenzidine
U.S. – United States
UV – Ultraviolet
VEGF-A – Vascular endothelial growth factor A
VH – Variable heavy
VL – Variable light

# 1. Introduction

## 1.1 Breast Cancer

Breast cancer is by far the most common cancer in women worldwide and the second most common cancer overall (Bray et al., 2013). Nearly 1.7 million new cases were diagnosed in 2012, which represents about 12% of all new cancer cases and 25% of all cancers in women. Over 522 000 women died in 2012 due to breast cancer, being also the first to second (in less developed and more developed countries, respectively) leading cause of cancer associated deaths in women and the fifth in both sexes (Ferlay, et al. 2013). Breast cancer incidence rates vary greatly worldwide, being higher in developed countries (Figure 1.1). However, almost 50% of breast cancer cases and 58% of deaths occur in less developed countries, where incidence is rising due to life expectancy and urbanization increase, and also due to adoption of western lifestyles (WHO, 2014). In Portugal, breast cancer is the cancer with the highest incidence rate; in 2012, over 6000 new cases were diagnosed and it caused over 1500 deaths (Ferlay, et al. 2013). Breast cancer survival rates vary greatly worldwide, ranging from 80% or over in North America, Sweden and Japan to around 60% in middle-income countries, and below 40% in low-income countries (Coleman et al., 2008). The low survival rates in less developed countries are mainly due to the lack of early detection programmes, which results in a high proportion of women that are diagnosed with late-stage disease, as well as by the lack of adequate treatment facilities (WHO, 2014).

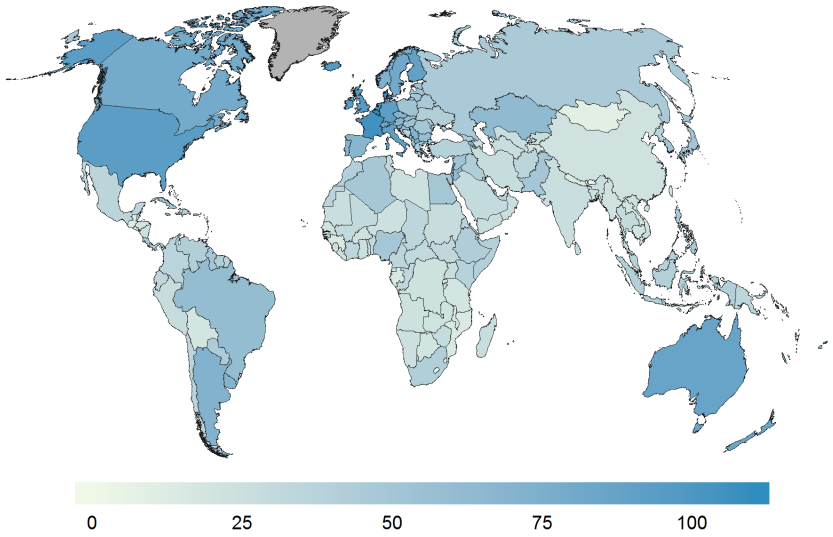


**Figure 1.1 - Estimated breast cancer incidence worldwide in 2012. Estimated age-standardized rates (world) per 100 000. Source: Ferlay et al, 2013.**

The most effective way to fight and control breast cancer is early detection, which considerably improves the disease outcome and survival rate (WHO, 2014). However, this does not always happen, as the above numbers indicate. Those reflect the impact of breast cancer in the world, which is increased by affecting an organ of great symbolism for femininity individually and socially. Considering all this, it is clear the need and importance of optimizing and discovering new possibilities for breast cancer therapy, which keeps on being a major area of research in medical oncology (Stopeck et al., 2012).

### 1.1.1 Breast Cancer Treatment: Targeted Therapy Relevance

For most patients with operable breast cancer, the primary therapy consists in tumor (sometimes along with regional lymph nodes) removal by surgery (Akram and Siddiqui, 2012; National Cancer Institute, 2009). Depending on the tumor characteristics and the stage of the disease several treatment options may be viable, and must be considered by the doctor (or specialist team in charge of the case) and the patient together (Cancer Research UK, 2014a). Also, neoadjuvant therapy may be needed in locally advanced breast cancer cases to reduce the tumor before surgery, so that it may result in more effective and less invasive outcome (National Cancer Institute, 2009).

After surgery, many strategies may be used to increase the chance of long-term disease-free survival, and these are specially indicated for patients who have a higher risk of breast cancer recurrence. Adjuvant therapy intends to kill or stop the growth of any cancer cells that remain in the patient's body and might not be detected. These kinds of therapies can include chemotherapy, hormonal therapy, targeted drugs (i.e. trastuzumab Herceptin®), radiation therapy, or a combination of treatments. Radiation therapy acts locally, and is aimed at cancer cells that may have been left in tissues in close proximity to the breast after surgery, such as those from the chest wall or nearby lymph nodes. On the other hand, systemic therapies can also reach and affect cells that may have spread to other locations (National Cancer Institute, 2009). Chemotherapy is a standard option that consists in the administration of cytotoxic drugs that interfere with cells with a high proliferating index, like cancer cells. However, since enhanced proliferation is also detected in epithelia and bone marrow cells, these drugs are highly toxic for the organism, resulting in the side effects commonly associated with chemotherapy (National Cancer Institute, 2014; Stopeck et al., 2012).

These potentially harmful and severe side effects should be taken into account before treatment, so that patients that are not likely expected to benefit from chemotherapy should not have to be subjected to it (Akram and Siddiqui, 2012). Nevertheless, chemotherapy is a conventional choice, once it is usually considered that if this treatment brings a small advantage to the patient even regarding the side effects, then this is better than none at all (Stopeck et al., 2012).

Although there are standard therapy procedures adopted in most breast cancer cases, it is long known and increasingly clear that, as for other cancers, there are molecularly distinct subtypes of breast cancers, that may require or take a greater benefit from different therapeutic approaches (Jackson and Chester, 2014).

Accordingly, breast cancer tumors are tested to check hormone receptors (estrogen receptor, ER, or progesterone receptor, PR) and human epidermal growth factor receptor 2 (HER2) status, which allows the identification of key molecular subtypes of human breast tumors and currently guide choice of therapy. For patients with hormone-receptor positive (ER- or PR-positive) breast cancer, hormone therapy with tamoxifen (drug that prevents estrogen from binding to the ER) and aromatase inhibitors (that block estrogen production by the body), or ovarian suppression (in premenopausal patients) represent the backbone of treatment, along with chemotherapy (Akram and Siddiqui, 2012). Since ER and PR are not exclusive to the neoplastic cells, hormone therapies can affect other tissues and organs that express them, with the possibility of causing some side effects, making hormone therapies 'semi-targeted' (Stopeck et al., 2012).

For patients with HER2-positive breast cancer, regardless of any other characteristics, the introduction of therapies directed against HER2 has dramatically improved prognosis (Akram and Siddiqui, 2012). HER2 (neu, C-erb B2) is a transmembrane receptor that belongs to the EGFR family of receptor tyrosine kinases. Contrarily to the other receptors of the HER family, no known natural ligand exists for HER2 (Mukai, 2010). HER2 can be overexpressed in breast cancer and in other solid tumors; HER2 gene amplification and protein overexpression is present in 15–30% of breast cancers and is directly linked to deregulated activation of the intracellular mitogenic signaling, leading to a more aggressive behavior and therefore poor prognosis (Akram and Siddiqui, 2012; Elloumi et al., 2012; Mukai, 2010; Stopeck et al., 2012) Anti-HER2 targeted agents are an example of molecularly targeted therapy, a type of treatment that blocks the growth and spread of cancer by interfering with specific molecules and pathways related to tumor progression and malignancy. These agents can act by inducing apoptosis of cancer cells, by blocking specific enzymes and growth factor receptors involved in cancer cell proliferation, cellular invasion and metastasis, or by modifying the function of proteins that regulate gene expression and other cellular functions. Since targeted therapy focuses on molecular abnormalities specific to cancer, it is expected to be similar, or even more effective, and less harmful to normal cells than systemic chemotherapy, which simply interferes with all rapidly dividing cells (Huang et al., 2014; Joo et al., 2013).

Trastuzumab (Herceptin®), an IgG1 humanized monoclonal antibody (mAb) that specifically binds to the extracellular domain of human HER2, has been a major success story in the field of targeted drugs (Mukai, 2010; Ribatti, 2014). Herceptin® was the first anti-HER2 targeted agent approved for clinical use in breast cancer patients (Mukai, 2010). This antibody (Ab) binds to the extracellular domain of HER2 receptor on the cells, and consequently induces the expression of anti-angiogenic factors, suppresses pro-angiogenic factors, and mediates Ab-dependent cytoxicity (Ribatti, 2014). Its clinical development relied upon selecting only those patients with the highest likelihood of response, or HER2 overexpressing tumors – “HER2 positive” (HER2+) – based on the principle that the target is also the best predictor of response. HER2+ classification is carefully attributed by measuring and scoring HER2 expression by standard clinical assays and accepted criteria. HER2 overexpression is thought to be a signal of the reliance of the cancer cell on HER2 signaling (“oncogene addiction”). Therefore, inhibition of HER2 signaling should result in a therapeutic effect mostly, if not exclusively, within cancer cells that are HER2+ and not be toxic to normal cells (Stopeck et al., 2012).

Despite its added value, 70% of HER2+ patients do not respond to treatment with Trastuzumab. In addition, acquired resistance to trastuzumab develops in a majority of patients with metastatic breast cancer, due to progression of the disease. This is mainly due to the development of mutations within the specific target-encoding gene, which can lead to a decrease in the affinity of the Ab towards it (Elloumi et al., 2012; Mukai, 2010). This problem has been managed by the introduction of other HER2-targeting agents with different mechanisms of action that can complement trastuzumab, and thus be an option in cases of resistance. Lapatinib (Tykerb®) is an oral dual tyrosine kinase inhibitor targeting HER1 and HER2 which blocks the downstream signaling pathways of these receptors (Mukai, 2010). Pertuzumab (Perjeta®) is also a mAb but this molecule targets a different

extracellular domain of HER2, specifically the dimerization region so that its binding inhibits HER2 dimerization with other HER receptors. This results in critical cell signaling inhibition and Ab-dependent cellular cytotoxicity activation. Furthermore, combined administration of pertuzumab and trastuzumab has demonstrated superior antitumor effects compared to either Ab alone in preclinical experiments (Joo et al., 2013; Stopeck et al., 2012). This leads to a second way of dealing with the limited efficacy of trastuzumab: combinatory therapy. This approach is becoming obligatory and these kinds of treatments are now being studied not only as monotherapies, but also in combination with other targeted therapies or with chemotherapy, with the goal of establishing the most effective combinations, with minimal side effects and a low damage to normal dividing cells (Elloumi et al, 2012; Joo et al., 2013; Stopeck et al., 2012). In this context, the Ab-cytotoxic drug conjugate ado-trastuzumab emtansine (Kadcyla®), was recently approved by the FDA (2013) (Joo et al., 2013). It consists in trastuzumab linked to the cytotoxic agent mertansine (DM1). DM1 inhibits tubulin polymerization, interfering with mitosis and promoting apoptosis. Kadcyla® combines the anti-HER2 activity of trastuzumab with the targeted intracellular delivery of DM1, with the advantage that its cytotoxic activity will be specifically carried out in HER2+ tumor cells (Ballantyne and Dhillon, 2013; Breastcancer.org, 2014a).

There are also some targeted therapies specific for other molecules that can be used in breast cancer; everolimus (Afinitor®) is indicated in advanced hormone receptor-positive, HER2-negative tumors; it is a transduction signal inhibitor targeting mTOR complex 1, consequently reducing the activity of its downstream effectors, which results in a decrease in cell proliferation, angiogenesis and glucose uptake (National Cancer Institute, 2014). Bevacizumab (Avastin®) is a recombinant humanized mAb that binds to vascular endothelial growth factor A (VEGF-A) and blocks angiogenesis (Elloumi et al., 2012). However, these are not consensually approved (Cancer Research UK, 2014b; Elloumi et al., 2012; Pazdur, 2013).

Clearly, the progress in targeted therapies for breast cancer has mainly been made on anti-HER2 drugs, which means other types of breast cancer are currently not benefiting from this type of treatment. Yet, breast cancer heterogeneity is not limited to the presence or absence of ER, PR and HER2 expression – tumors that are negative for all these three conditions are called triple negative. These correspond to 15 to 20% of breast cancers, and are more often high grade, highly proliferative, and locally advanced at diagnosis. However, triple negative cancers are not a single type of disease, harboring distinct phenotypic and functional characteristics resulting from genetic diversities and non-genetic influences (Huang et al., 2014; Stopeck et al., 2012). Consequently, no single therapy is likely to be effective for all patients and current treatment options are limited, so their management is challenging (Akram and Siddiqui, 2012; Stopeck et al., 2012). Although a subset of types of cancer may respond to chemotherapy, advantage may come from the identification of new specific pathways and targets for therapy (Akram and Siddiqui, 2012; Joo et al., 2013). Even if triple negative cases are those who have more to gain with the discovery of new targets for therapy, the remaining could also benefit. Surely, some ER, PR and HER2 positive cancers may express other molecules that are targeting candidates, given that the oncogenic process implies the deregulation of several proteins and signal transduction pathways at the same time. Despite the several options of treatment available,

there are always cases of failure, and new therapeutic options targeting other molecules would be welcome and allow improving the chances of success, when patients are appropriately selected accordingly to the alternative predictive biomarkers they express (Jackson and Chester, 2014).

## 1.2 The Notch Signaling Pathway

Notch signaling is an evolutionarily conserved developmental pathway, expressed in a wide range of tissues and organisms, indicating its multiplicity of functions (Calaf et al., 2014; Izrailit and Reedijk, 2012; Takebe et al., 2014). During development, it is critical for cell fate determination and homeostasis, by controlling cell growth and proliferation, migration, differentiation and apoptosis, and participating in tissue patterning, morphogenesis, organogenesis and angiogenesis (Aste-Amézaga et al., 2010; Calaf et al., 2014; Han et al., 2011; Takebe et al., 2014). Although classically known for its vital role in embryogenesis, it is also implicated in postnatal hematopoiesis, breast development, gastrointestinal epithelial maturation, immune regulation, vascular development and somatic stem cell maintenance, renewal and differentiation (Sethi et al., 2011; Takebe et al., 2014). Given this diversity, Notch signaling outcome largely depends on the cell-context, microenvironment and crosstalk with other signaling pathways (Takebe et al., 2014).

Different Notch receptors have different functions and regulate cell fates differently. In mammals, they are four (Notch1, 2, 3 and 4) and each of them have five ligands (Jagged 1 – Jag1 and Jagged2 – Jag2, and Delta-like 1 – DLL1, 3 – DLL3 and 4 – DLL4) (Sharma et al., 2012; Takebe et al., 2014). All Notch receptors and their ligands are single-pass transmembrane proteins, so the Notch signaling is activated by direct cell–cell contact, consisting of a short-range intercellular communication system between neighboring cells (Izrailit and Reedijk, 2012; Suman et al., 2013; Takebe et al., 2014).

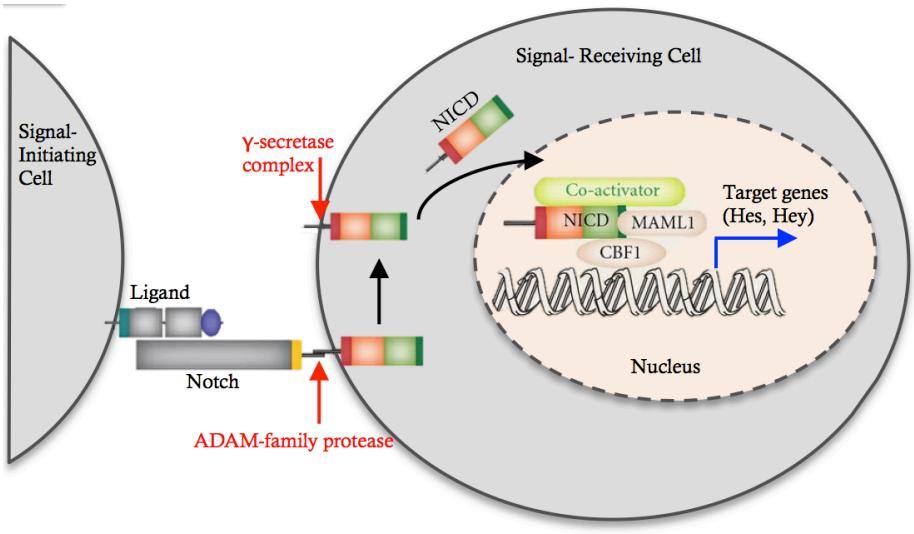


**Figure 1.2 - Notch canonical signaling pathway overview. Red arrows – proteolytic cleavages. Adapted from Ersvaer et al., 2011.**

The Notch receptors and their ligands contain multiple epidermal growth factor (EGF)-like repeats in the extracellular region required for their binding. In the receptors, follows a negative regulatory domain (NRR) that, in the absence of ligand, maintains the receptor in a protease-resistant

conformation (Aste-Amézaga et al., 2010; Takebe et al., 2014). Ligand binding activates the pathway by triggering a conformational change in the receptor NRR that is followed by two sequential proteolytic cleavages; the first is in the extracellular domain and is mediated by metalloproteases of the ADAM (A Disintegrin And Metalloproteinase) family, and the second occurs within the transmembrane domain and is mediated by the γ-secretase complex. This results in the release and translocation of the active Notch intracellular domain (NICD) into the nucleus, where it interacts with the DNA-binding protein CSL (C-promoter binding factor 1, CBF1 in humans), converting it from a transcriptional repressor to activator by recruiting cofactors such as Mastermind-like proteins. This activates the transcription of Notch target genes, like HES1, HES5, HEY2, NRARP, DELTEX1, and c-MYC (Figure 1.2) (Aste-Amézaga et al., 2010; Izrailit and Reedijk, 2012; Takebe et al., 2014).

### 1.2.1 Notch Signaling Pathway and Breast Cancer

Deregulated Notch signaling is associated with numerous human diseases, including a broad spectrum of cancers. It is generally physiologically involved in the development and maintenance of normal tissues that originate and are recapitulated in those different forms of cancer. Deregulation of the pathway might affect normal cell differentiation and metabolism, cell cycle progression, angiogenesis and possibly self-renewal and immune function (Bolós et al., 2013; Calaf et al., 2014). Moreover, some of the Notch target genes are well known to have significant roles in carcinogenesis and tumor progression (Takebe et al., 2014).

An oncogenic role for Notch was first discovered in T-cell lymphoblastic leukemias (T-ALL), in which Notch-1 gain-of-function mutations are found in over 50% of cases. Since then, aberrant Notch signaling was subsequently identified in many solid tumors, including pancreatic, prostate, renal, lung, glioblastoma multiforme, sarcomas, gastroenteropancreatic neuroendocrine tumors, cervical, melanoma, head and neck, and breast cancers. In those, however, current evidence suggests that aberrant Notch activation occurs through transcriptional and post-translational mechanisms resulting in ligand/receptor abundance rather than activating mutations or amplification of the Notch *loci*. Ligands mutations are also infrequent (Izrailit and Reedijk, 2012; Takebe et al., 2014).

High expression of Notch receptors and ligands has been recently implicated in breast cancer pathogenesis and associated to poor clinical outcomes. Reedjik et al (2005) identified high levels of Jag1 or Notch1 in about 25% of human breast tumors by *in situ* hybridization, and associated high levels of co-expression of Notch1 and Jag1 with reduced overall survival. Immunohistochemistry studies revealed increased expression of several Notch receptors and ligands in human breast cancer specimens compared to normal breast tissue, and different co-expression patterns (Rizzo et al., 2008; Mittal et al., 2009). Notch1 down-regulation and inhibition by several methods in human breast cancer cell lines has demonstrated anti-tumor effects, like inhibition of invasion and metastasis (Wang et al., 2011), proliferation inhibition and apoptosis induction (Sharma et al., 2012), mammosphere formation inhibition and size reduction (Simmons et al., 2012; Suman et al., 2013). Mammary tumor regression and disease recurrence prevention were also demonstrated in transgenic mice models (Simmons et al., 2012).

Notch association to worse prognosis has also been exploited. As in the case of normal stem cells, Notch signaling plays an important role in maintenance of cancer stem cells (CSCs), capable of

self-renewal and differentiation in malignant cells (Joo et al., 2013; Sharma et al., 2012). The CSCs are believed to constitute a subpopulation in various cancers that is responsible for tumor initiation, progression, metastasis, and also resistance to therapy, being therefore the source of recurrence (Joo et al., 2013; Suman et al., 2013; Takebe et al., 2014). Notch activity was identified as a marker for cells with stem cell-like properties, and there is evidence that it is required for breast CSCs renewal and expansion, more specifically, that Notch4 has a role in their maintenance and tumor initiation, while Notch1 participates in tumor growth and proliferation (Izrailit and Reedijk, 2012; Takebe et al., 2014). Furthermore, Notch1 inhibition sensitizes breast cancer cells to chemotherapy (Suman et al., 2013).

Another connection between Notch signaling and poor clinical outcomes is the correlation between ligand Jag1 expression in breast cancer cells and bone metastasis, which correspond to 70% of metastatic breast cancer. Sethi et al (2011) identified Jag1 as a clinically and functionally important mediator of bone metastasis by activating the Notch pathway in bone cells.

Considering the above mentioned, it is clear that Notch signaling provides a potential therapeutic target for breast cancer treatment.

### 1.2.2 Notch-Targeting Therapeutic Agents

Different strategies are being developed to block Notch signaling for therapeutic purposes. One approach is to prevent the proteolytic cleavages that result in the release of NICD by treatment with γ-secretase inhibitors (GSIs) (Aste-Amézaga et al., 2010; Sharma et al., 2012). GSI treatment of Notch-induced breast cancer cells *in vitro* reduces Notch activity and tumor proliferation, and inhibits metastasis of mammary tumors (Groth and Fortini, 2012; Izrailit and Reedijk, 2012). However, this method has poor selectivity, interfering with the proteolysis of all four Notch receptors as well as with multiple additional proteins involved in other signaling pathways. This leads to significant limitations to their therapeutic potential, namely severe adverse effects resulting from GSI-induced toxicity and dysfunction of several systems (Aste-Amézaga et al., 2010; Izrailit and Reedijk, 2012; Sharma et al., 2012).

The second most explored class of agents under development is mAbs, which are expected to overcome GSIs restraints (Takebe et al., 2014). mAbs have the advantage of specificity, but also versatility, once they can be developed against any relevant member of the pathway. It has already been demonstrated that mAbs developed against Notch1 specifically bind to and inhibit Notch1 signaling (Aste-Amézaga et al., 2010), inhibit Notch1-driven tumors growth, and have reduced side effects (Wu et al., 2010). In fact, several Notch-targeting mAbs are currently in early phase clinical development, including mAbs against Notch1, Notch2/3 and DLL4, whose effects are being studied in subjects with solid tumors (Takebe et al., 2014).

## 1.3 Antibodies

Antibodies (Abs) belong to a family of globular proteins called immunoglobulins (Igs) and are the antigen-binding proteins present on the B-cell membrane and secreted by plasma cells (Elgert, 2009; Goldsby et al., 2003). They have a fundamental role in the humoral response of the immune system, having functions like binding specifically to the pathogen or its products that elicit the immune

response, and recruiting other cells and molecules to destroy the pathogen once the antibody has bound (Murphy, 2012).

### 1.3.1 Antibodies Structure

All Abs share a common structure (Figure 1.3 A) of four peptide chains that consist of two identical light chains (LC), of about 25 kDa each, and two identical heavy chains (HC), larger polypeptides of 50 kDa or more each. The two HC are linked to each other by disulfide bonds, and each HC is linked to a LC by a disulfide bond (Goldsby et al., 2003; Murphy, 2012). There are two types of LC, lambda (λ) and kappa (κ). A single Ab molecule contains only one LC type, never both, and no functional difference has been found between Abs having each of them (Murphy, 2012).

The class and thus the effector function of an Ab is defined by the structure of its HC, more specifically by its carboxy (C)-terminal part or constant region. There are five main HC classes or isotypes, some of which have several subtypes, and these determine the functional activity of an Ab molecule. The five major classes of Igs are IgG (γ), IgM (μ), IgA (α), IgE (ε) and IgD (δ). IgG is the most abundant Ig. Each class can have either a λ or κ LC (Goldsby et al., 2003; Murphy, 2012).

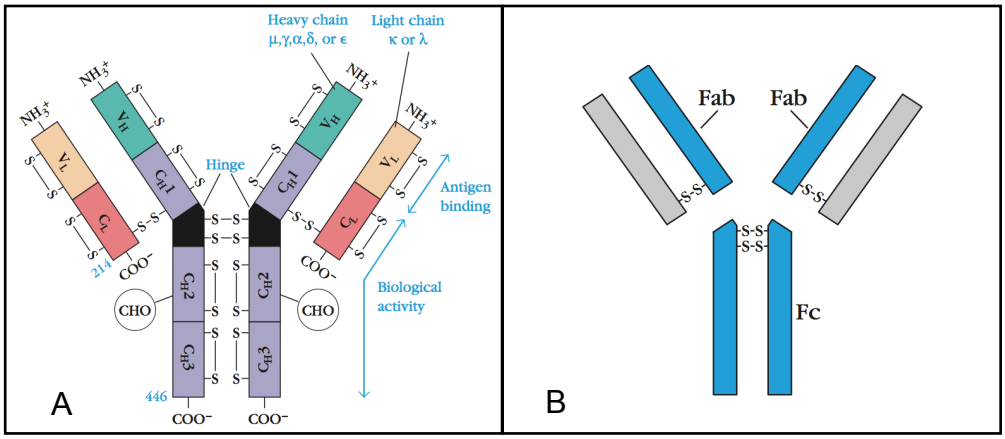


**Figure 1.3 – The Ab molecule. A – Schematic diagram of Ig structure. B – Fab and Fc fragments; Grey – LC; Blue – HC. Adapted from Goldsby et al., 2003.**

The first 100-113 amino acids of the amino (N)-terminal region of a LC or HC greatly vary between Abs of different specificity. These segments of highly variable sequence are called V regions: VL in LC and VH in HC (Elgert, 2009; Goldsby et al., 2003). Since the amino acid sequence dictates the three-dimensional structure of the protein, the unique sequence of amino acid residues for each V region leads to a large diversity of structures, which accounts for antibody specificity (Elgert, 2009). In contrast, the remaining portion of the molecule has a relatively constant sequence within the same Ab class, making up the C regions (CL in light chains and CH in heavy). The HC C region contains multiple domains, either three or four, depending on the antibody class, and those are numbered from the N-terminal end to the C-terminus CH1, CH2, CH3, and CH4 (Goldsby et al., 2003).

The LC and HC association forms the Ab molecule, which comprises three equal-sized globular portions of 50 kDa joined by a flexible polypeptide chain, the hinge region. The final result is a Y-shaped structure; the two upper arms are identical and each is formed by the association of a LC with the amino-terminal half of the HC (Fd fragment). More specifically, the VL region interacts non-covalently with the VH region, and the CL region with the CH1 region. These are called Fab fragments

("fragment, antigen binding"), and contain the antigen binding activity. The other fragment is the Fc fragment ("fragment crystallizable"), which corresponds to the remaining constant HC domains, and is the part of the Ab molecule that interacts with effector molecules and cells (Figure 1.3 B) (Elgert, 2009; Murphy, 2012).

Looking closer to the V regions organization, there are three particularly variable segments that can be identified in both VH and VL domains, designated as hypervariable regions or complementarity-determining regions (CDRs) – CDR1, CDR2, and CDR3. In the HC they are located roughly at residues 30 to 36, 49 to 65 and 95 to 103, respectively, and in the LC at residues 28 to 35, 49 to 59 and 92 to 103 (Murphy, 2012). The regions between the CDRs, which comprise the rest of the V domain (about 85%), show less variability and are termed the framework regions (FR). There are four in each V domain, FR1, FR2, FR3 and FR4, and they are responsible for the CDRs positioning on the surface of the chain, when the VH and VL domains are joined. This creates a single hypervariable region at the end of each Fab. Due to the fact that the amino acid sequences and length of the CDRs determine the shape and ionic properties of the antigen-binding site, the CDRs define the specificity of the antibody (Elgert, 2009; Murphy, 2012).

The Abs diversity is generated by somatic recombination in B lymphocytes, whereby V-gene segments are linked to other gene segments. The LC families contain VL, JL, and CL gene segments and the HC family contains VH, DH, JH, and CH gene segments. C gene segments encode the constant regions (Goldsby et al., 2003). In the final assembled LC V domain, the VL gene segment encodes FR1 to 3, CDR1 and 2, and two thirds of CDR3, while JL encodes the rest of CDR3 and FR4. In the HC V domain, VH gene segment encodes FR1 to 3 and CDR1 and 2 and JH encodes FR4. HC CDR3 is created *de novo* in developing B cells by the joining process, containing the entire DH as well as portions of VH and JH gene segments (Figure 1.4) (Paul, 2013).

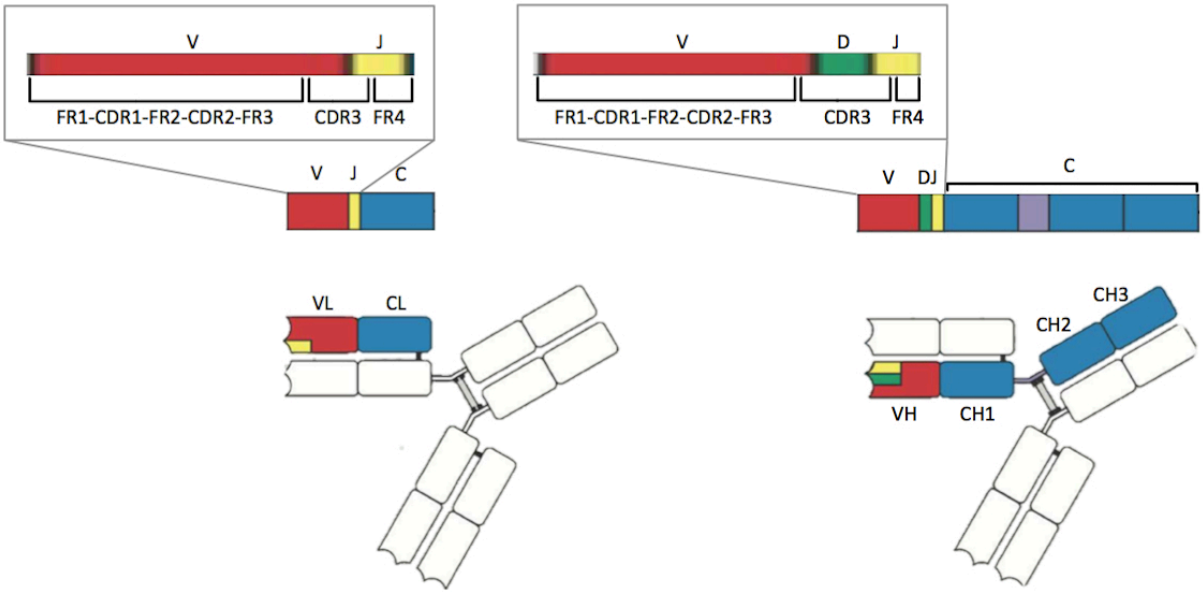


**Figure 1.4 - Final assembly of the Ab gene segments. Variable chain gene segments contribution to the antigen-binding site. Adapted from Murphy, 2012.**

Additional diversity is generated during recombination by imprecise joining, especially at the VH-DH and DH-JH junctions, which allows nucleotides loss and gain. Altogether, this results in the

creation of nearly random HC CDR3 sequences, where the greatest diversity of an Ab repertoire is focused (Burton, 2001; Paul, 2013).

Although the undeniably fundamental role of the V regions, CH1 and CL domains may also contribute to Ab diversity by allowing more random associations between VH and VL domains than would occur if these were driven by the VH/VL interaction alone. The presence of CH1 and CL appears to increase the number of stable VH and VL interactions that are possible, thus contributing to the overall diversity of Ab molecules that can be expressed by an animal. These considerations have important implications for building a diverse Ab repertoire (Goldsby et al., 2003).

### 1.3.2 Therapeutic Antibodies

Abs are important tools for research and medicine, being monoclonal antibodies (mAbs) the most useful (Goldsby et al., 2003). mAbs are monospecific Abs made by identical immune cells that derive from a unique clone and thus are specific for a single epitope (Goldsby et al., 2003; Xin et al., 2013).

Initially, in the clinical medicine field mAbs were used primarily as *in vitro* diagnostic reagents (Goldsby et al., 2003), but soon their value as therapy agents was noticed. Presently, Ab-based therapeutics is an important component in the treatment of an increasing number of human malignancies, including cancer, infectious diseases, transplantation, allergy, asthma, and some autoimmune diseases. The mechanism of action of a mAb includes neutralization of substances, blocking of receptors, binding to cells and modulating the host immune system by inducing effector functions, such as Ab-dependent cellular cytotoxicity, or a combination of these effects (Brissette et al., 2006; Schirrmann et al., 2011 ). These molecules can also be used as drug delivery systems or Ab-cytotoxic drug conjugates, specifically inducing death to target cells that engulf the conjugates after they bind (Joo et al., 2013).

Abs are one key component of molecularly targeted therapy, perfectly fitting its demands given their target specificity and low toxicity, in addition to well-defined pharmacological properties, such as high affinity and long serum half-life (Brissette et al., 2006; Dantas-Barbosa et al., 2012).

There are currently more than 250 therapeutic mAbs undergoing clinical trials (Ribatti, 2014). To date, a total of 39 therapeutic mAbs have already been approved by the U.S. Food and Drug Administration (FDA) and the European Medicines Agency (EMA), 15 of which are directed to cancer treatment, and 4 are under review for approval (2 of them also for cancer) (Landes Biosciences, 2014). Furthermore, Abs comprise the second-largest category of biological medicines in clinical development (Dantas-Barbosa et al., 2012). Among the mAbs approved for cancer treatment are Trastuzumab, Pertuzumab and Trastuzumab emtansine, all for HER2+ breast cancer (see section 1.1.1).

Even though everything points to the successful establishment of Abs in therapeutics, the Ab development method chosen may limit the possibility of overcoming all challenges from idealization to approval. Among them are relatively long development times, high production costs, difficult manufacturing, poor effector functions and pharmacokinetics but, most importantly, safety concerns related to immunogenicity, when these are obtained by animal immunization and hybridoma technology (Brissette et al., 2006; Dantas-Barbosa et al., 2012; Ribatti, 2014; Schirrmann et al., 2011).

The latest problem has been solved by chimerization and humanization of mouse Abs using genetic engineering techniques, which are still popular alternatives. Chimeric Abs consist of a combination between the constant domain of human IgG and the variable antigen-specific mouse domain (30 to 35% mouse proteins and 60 to 65% human proteins). Humanized mAbs are originated by the implantation of the mouse CDR in the framework of human IgG, and thus the human part of the Abs is between 90 to 95% (Elloumi et al., 2012). However, it is already possible to produce fully human Abs utilizing different approaches, such as phage display technology or genetically engineered mice (or other animals), which offer more advantages because they are less antigenic, better tolerated and have a longer circulation time (Dantas-Barbosa et al., 2012; Ribatti, 2014; Xin et al., 2013). In fact, of the 39 approved therapeutic mAbs, 11 are human, 4 derived from phage display technology and 7 obtained by genetically engineered mice (Landes Bioscience, 2014). The latest were generated by replacement of animals' native Ab genes by pure human homologues. Upon immunization, these animals develop human IgG in response to the antigen, which then can be produced by conventional hybridoma technology (Dübel, 2007; Ribatti, 2014). This implies that, despite solving the immunogenicity issue, this option is not free of the remaining hybridoma technology constraints; although well established, this technology is laborious, and it is limited by the possible instability of the aneuploid cell lines and by the animal immune system inability to provide high-affinity Ab against toxic or highly conserved mammal proteins (Dantas-Barbosa et al., 2012; Schirrmann et al., 2011).

On the other hand, the phage display technology is completely independent from any immune system and uses a completely *in vitro* selection process, offering a rapid approach for generating and identifying highly specific reagents relatively quickly, with a resulting decrease in both development time and costs (Brissette et al., 2006; Schirrmann et al., 2011). Also, the first fully human therapeutic Ab to be approved was isolated by this method – the anti-TNF-α Ab adalimumab (Humira®) for rheumatoid arthritis treatment (Schirrmann et al., 2011).

## 1.4 Phage Display Technology

The phage display concept was first introduced by George Smith in 1985, when he demonstrated that foreign DNA fragments could be fused to a coat protein-coding gene of a non-lytic filamentous phage and expressed as a fusion protein on the virion surface without disturbing the infectivity of the phage (Ahmad et al., 2012). Five years later, McCafferty et al showed that cloned Ab fragments could be similarly displayed on phage particles as functional proteins that retained an active antigen-binding domain capability. Soon the first Ab gene repertoires or "libraries" had been generated for filamentous phage display and Ab selection (Ahmad et al., 2012; Schofield et al., 2014). Phage display is now a mainstream Ab and protein engineering platform, being one of its most successful applications the genetic selection of mAbs using large phage Ab libraries (Hoogenboom, 2005; Schofield et al., 2014).

The ultimate aim of phage display is the selection of peptides or proteins that can bind to a target of interest with high affinity, from a huge number of nonspecific candidates (Ahmad et al., 2012). This is achieved by using large combinatorial repertoires or libraries. Phage display libraries are constructed by mass cloning of a pool of genes encoding millions of variants of certain ligands into an adequate vector (like the phage genome), without need of individual identification of each gene,

allowing the effortless construction of diverse and rich collections (Ebrahimizadeh and Rajabibazl, 2014; Hoogenboom et al., 1998). This results in the fusion of those genes with a gene encoding one of the phage coat proteins that is carried by the vector. Upon expression, the coat protein fusion will be incorporated into new phage particles that are assembled in the bacterium (Hoogenboom et al., 1998). The fusion product is displayed on the phage surface, and can be selected for its specific binding properties (Hoogenboom et al., 1998; Schofield et al., 2014). Moreover, the genetic information encoding the displayed protein is contained within the phage particle, providing a direct link between phenotype and genotype, which is the fundamental aspect of phage display (Figure 1.5) (Carmen and Jermutus, 2002; Hoogenboom et al., 1998). This linkage allows the enrichment of the selected phages but also further manipulation and development of genetic studies of the correspondent displayed proteins, due to the possibility of recovering the genetic information (Brissette et al., 2006; Hammers and Stanley, 2014).

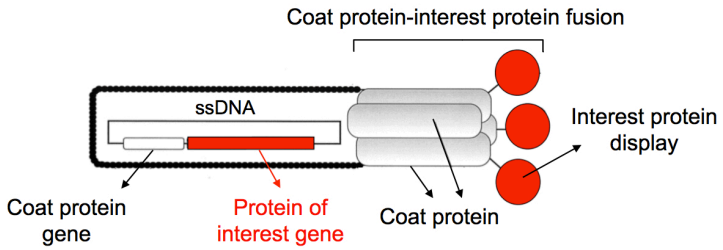


**Figure 1.5 - Generic schematic representation of a protein-displaying phage particle. Adapted from Yamabai, 2014.**

Phage display technology is robust, simple to use, highly versatile, easy to implement and inexpensive (Dantas-Barbosa et al., 2012; Hoogenboom et al., 1998). It can be exploited in broad spectrum of biomedical and pharmaceutical researches, including: identification of cell/tissue or disease-specific biomarkers; protein-protein interaction; receptor-ligand characterization; epitope mapping; gene delivery/targeting; and proteomic and functional genomic approaches (Tohidkia et al., 2012). Last but not least, it plays an important role in drug and vaccine discovery and development, which includes production of mAbs, with a number of molecules approved by FDA and many others in clinical trials (Dantas-Barbosa et al., 2012; Tohidkia et al., 2012)

### 1.4.1 Phage Panning

One of the advantages of phage display is the rapid identification and enrichment of target-specific binders from a large excess of non-binding clones, through a process known as biopanning. This is achieved by multiple rounds of phage incubation with a chosen target, washing to remove non-specific or low affinity phages, elution to retrieve specifically bound phages and their amplification (Figure 1.6) (Hoogenboom et al., 1998). In the first round, the selected subpopulation should be a small fraction of the very large initial library, so it will be amplified in the appropriate host cells in order to achieve a significantly higher representation of each individual phage in the amplified stock (Smith and Petrenko, 1997). A single round of selection can enrich for a specific phage 20 to 1000-fold. Performing multiple selection rounds with selected (eluted and amplified) phages increases the enrichment of specific binders (Watkins and Ouwehand, 2000), allowing the isolation of potentially

very rare antigen-binding clones (Hammers and Stanley, 2014). The panning steps should be repeated until the identification of optimal binders, which usually takes 3 to 5 rounds (Dantas-Barbosa et al., 2012).

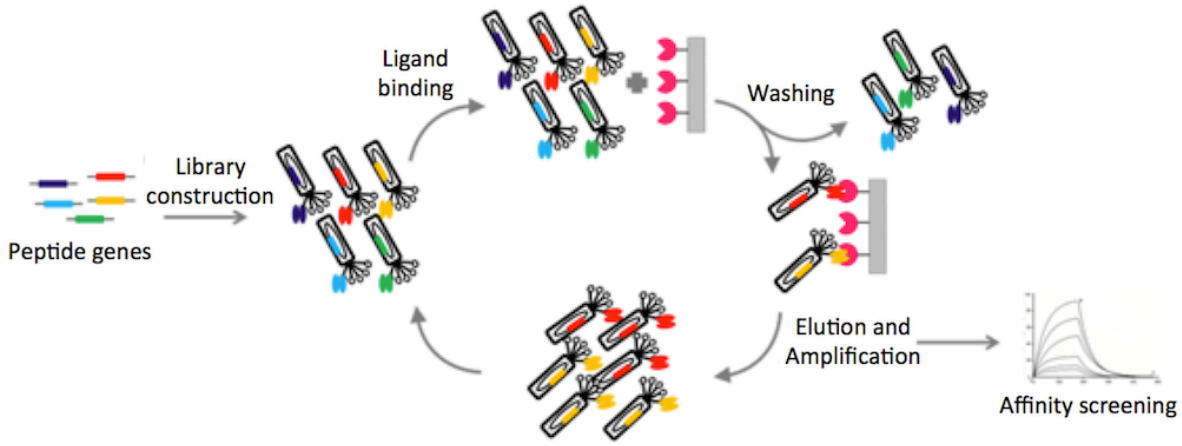


**Figure 1.6 - Phage display technology overview. Adapted from Biotem, 2011.**

The success of ligand phage display hinges on the combination of this display and enrichment method, with the synthesis of large combinatorial repertoires on phage (Hoogenboom et al., 1998). Various types of selection strategies can be used. The simplest and most common approach is to use a purified antigen immobilized on a solid support, which requires a good source of antigen (Watkins and Ouwehand, 2000), but affinity selection can be performed by using low antigen concentrations. Also, if the selection procedure is appropriately designed, antibodies can also be selected on the basis of kinetic properties, improved specificity, or phage infectivity (Ahmad et al., 2012). Other systems may be required when pure antigen is not available (Ahmad et al., 2012), or when the aim is to obtain Abs against unknown (e.g., novel markers studies) and non-immunogenic cell surface antigens in their native form (Dantas-Barbosa et al., 2012). These include the use of eukaryotic cells and tissue fragments as targets. Morover, *in vivo* and double recognition panning selections to select for bispecific mAbs have also been reported (Dantas-Barbosa et al., 2012; Hammers and Stanley, 2014). As the target antigen is presented in a more complex environment, such as cell surface, the selection procedure becomes more difficult, once many other antigens may be present and the expression levels may not be constant (Dantas-Barbosa et al., 2012), requiring more rounds of panning and more sophisticated protocols (Hammers and Stanley, 2014).

### 1.4.2 Antibody Libraries

Ab phage technology has revolutionized the way mAbs are produced, currently being the most widespread method for the display and selection of large collections of Abs and for the engineering of selected Abs (Hoogenboom, 2005; Sommavilla et al., 2010). As mentioned before, this technique has the advantage of allowing the direct isolation of fully human Abs (Dantas-Barbosa et al., 2012), which can subsequently be manufactured using fully *in vitro* processes, offering greater flexibility during their production and greater opportunities for optimization after their creation. Ab production and screening is cheaper, easier and faster by phage display, and it allows further manipulation of selected binders

due to the gene sequence availability (Hoogenboom et al., 1998; Shukra et al., 2014). Furthermore, it is a highly versatile technology (Hoogenboom et al., 1998).

#### 1.4.2.1 Formats for Antibody Display

*E. coli* production or phage display of complete IgG molecules is problematic due to limitations of the prokaryotic folding machinery, so it has rarely been described (Schirrmann et al., 2011; Schofield et al., 2014). Therefore, smaller Ab fragments are used, and the most common are Fab fragments and single chain variable domains (scFv), which consist of only one polypeptide chain, composed of the VH and the VL Ab regions joined by a short peptide linker (Ahmad et al., 2012; Schofield et al., 2014). Although it shows up as a need, some authors believe the IgG phage display cannot compete with Fab or scFv display and the additional Fc part is not a benefit for the display of binders in this prokaryotic system (Schirrmann et al., 2011).

The choice of which format to use should take into account the library's final application. The smaller size of the scFv format makes these libraries genetically more stable and of simpler construction than Fab libraries (Burton, 2001). Also, the expression of these smaller fragments has a less toxic effect on the cell, so scFv's are expressed at higher levels, increasing the library diversity (Carmen and Jermutus, 2002). On the disadvantage side, scFv's have the tendency to form dimers and higher order multimers in a clone-dependent and relatively unpredictable way (de Haard et al., 1999). This can be valuable in a certain context, namely if an enhanced binding avidity, which is the overall strength of the multiple binding interactions that may occur in a multivalent Ab-antigen complex (Goldsby et al., 2003), is useful to facilitate selection against some antigens. However, it usually complicates Ab selection and characterization, and the multimers avidity effect may lead to selection of fragments with very low affinity (Burton, 2001; Hoogenboom et al., 1998), which is the strength of interaction between a single epitope and a single antigen-binding site on an Ab (Goldsby et al., 2003).

The use of Fab fragments avoids these inconveniences. In Fab selections, affinity dominates over avidity, both because of the lack of multimerization and lower display frequency. Furthermore, a Fab can be converted into a whole Ab with a predictable maintenance of, or increase in, antigen affinity, which is unlikely for scFv's (Burton, 2001).

#### 1.4.2.2 Types of Antibody Libraries

There are several types of libraries. Each type has its limitations and is best suitable for different purposes. They are categorized based on the source of Ab genes used for their construction (Ahmad et al., 2012).

"Single-pot" libraries are called this way because they are designed to isolate Ab fragments against all or, at least, a wide variety of antigens. Once they are comprised of Ab fragments from a source of genes that is not biased to a specific antigen, they may be considered universal (Hoogenboom, 2005; Schirrmann et al., 2011). Nevertheless, they allow isolation of high-affinity Abs when very large repertoires are used. These libraries are particularly useful for the selection of human Abs to self, non-immunogenic or toxic antigens, which may be difficult to obtain with techniques that require immunization (Bazan et al., 2012; Hoogenboom et al., 1998). There are large pre-made non-immune repertoires available that suit most applications and are extensively used in industry and for

academic purposes (Hoogenboom et al., 1998; Hoogenboom, 2005). “Single-pot” libraries include naïve, semi-synthetic and synthetic libraries (Hoogenboom et al., 1998).

Naïve libraries are composed of natural rearranged V-genes from the IgM mRNA of B cells of non-immunized donors, isolated from diverse human or animal lymphoid sources. IgM Abs constitute the primary (unselected) Ab repertoire of an organism, and therefore their use reduces antigen-induced biases, being more diverse. The affinity of Abs selected from a naïve library is proportional to the size of the library (Hoogenboom et al., 1998; Hoogenboom, 2005). The main disadvantages of the Naïve libraries are the unknown history of the B-cell donors, the tendency to achieve increased cross-reactivity, and potentially limited diversity of the IgM repertoire (Bazan et al., 2012). However, random combination of VH and VL chains during the library construction will improve its diversity (Hoogenboom et al., 1998).

Semi-synthetic libraries have combinations of natural and synthetic diversity in order to increase natural diversity (Hoogenboom, 2005). The most common approach is to introduce synthetic randomized CDRs into germline V-gene segments or rearranged V-gene. This is usually carried out within the CDR3 regions, since they are the most diverse and essentially responsible for antigen binding, specially the HC CDR3 (Ahmad et al., 2012; Hoogenboom et al., 1998). Randomization is done by PCR or oligonucleotide directed mutagenesis and it may also involve other or all the CDR regions, and even FR regions (Bazan et al., 2012; Hoogenboom et al., 1998; Hoogenboom, 2005).

Synthetic libraries are constructed entirely *in vitro*, using fully synthetic sources. A certain number or, sometimes, a single artificially designed Ab framework is used as scaffold for the CDRs, which are all randomized, varying in composition and length (Ahmad et al., 2012; Schirrmann et al., 2011 ). The framework may be based on a limited number of generic or consensus V-gene sequences that are assembled *in vitro* (Bazan et al., 2012). Synthetic diversity bypasses the natural biases and redundancies of Ab repertoires created *in vivo* and allows control over the genetic makeup of V genes and the introduction of diversity (Hoogenboom, 2005). Synthetic libraries are of particular importance in isolating Abs recognizing self-proteins, once these specificities are avoided by the mechanisms of self-tolerance and clonal deletion *in vivo* (Watkins and Ouwehand, 2000).

Immune libraries are constructed with V-gene pools from the IgG mRNA of B cells from diverse lymphoid sources of immunized donors, i.e., which have an active immune response against a chosen antigen (Watkins and Ouwehand, 2000). An immune Ab repertoire will be enriched in Abs specific for that antigen, some of which will have been affinity matured by the host’s immune system, which allows achieving high-affinity Abs even when the library is relatively small ($10^7$ clones) (Bazan et al., 2012; Hoogenboom et al., 1998). The main disadvantage of immune libraries is that new Ab libraries have to be generated for every desired antigen making this a very laborious process (Watkins and Ouwehand, 2000). The immunization of animals is time consuming for itself, prediction of the immune response is not always possible, and it may be ineffective if self or toxic antigens are used (Bazan et al., 2012; Hoogenboom et al., 1998). Moreover, V gene combinatorial library construction is complex when both immune VH and VL genes are used (Watkins and Ouwehand, 2000). On the other hand, the availability of human immune libraries is limited due to ethical reasons; however, naturally immunized humans (by a infection or disease) may be used as donors, and a human library is the

obvious choice if the Abs are required for therapeutic purposes (Burton, 2001; Hoogenboom, 2005; Schirrmann et al., 2011). Nevertheless, immune libraries are valuable resources, typically generated and used in medical research to produce high quality and high affinity antigen-specific Abs that provide good analytical tools or when a range of different Abs for the same molecule is the goal (Bazan et al., 2012; Burton, 2001; Schirrmann et al., 2011). Further, human immune or disease-associated Ab libraries have identified Abs with very interesting properties unlikely to be present in single-pot libraries, and rare Ab specificities can be enriched by *in vitro* selection. These libraries also facilitate the investigation of the humoral immune system at a molecular level (Hoogenboom, 2005). Finally, although an immune repertoire is generally restricted to generating Abs against the immunization antigen, it may contain a significant number of unimmunized clones and, if sufficiently large, it can be utilized similarly to a naïve library (Bazan et al., 2012).

#### 1.4.2.3 Combinatorial Antibody Libraries Construction

The standard methodology for construction of natural source combinatorial phage Ab libraries consists in several straightforward steps. cDNA pools are produced from the mRNA extracted from the lymphoid source chosen, and used as template to amplify the VH and VL gene segments of a given Ig repertoire, by using specific primers that anneal to conserved regions of the V-gene families (Tohidkia et al., 2012). VH and VL chains may be separately amplified, and then randomly assembled by PCR to form the final Ab fragment. Those may also be randomly combined through sequential cloning of the VH and VL gene repertoires in the same vector (Frenzel et al., 2014; Watkins and Ouwehand, 2000). These and other strategies, like using several donors for the construction of a single library, will increase the original repertoire diversity, one of the most significant characteristics of phage Ab libraries. Besides diversity, another important feature of an Ab library is the size, and both will define the affinity and specificity spectrum of the Abs a library can produce (Tohidkia et al., 2012).

Library construction methods may differ in terms of origin of V-gene repertoires (immune, naïve, and synthetic library); number of donors; source of B-cells' sample; and molecular methods for amplifying and engineering the V-gene repertoires (e.g., RNA extraction procedure, primer sets for synthesis of cDNA and the VH and VL genes, the V-genes assembly procedure and vectors) (Tohidkia et al., 2012).

The key for a successful generation of Abs is the library used for the selection, and its performance depends both on technical aspects (such as library size and quality of cloning) and design features (which influence the percentage of functional clones in the library and their ability to be used for practical applications). Therefore, those should be carefully outlined (Frenzel et al., 2014).

### 1.4.3 Phage Display Vectors

A bacteriophage (phage) is a virus that specifically infects bacteria and replicates within it, and this explains why phages are commonly used as vectors for biotechnology and recombinant DNA research (Smith and Petrenko, 1997). Phage display, as its name indicates, is a technology that uses bacteriophages to connect proteins with the genetic information that encodes them, and allows the study of protein-protein, protein-peptide, and protein-DNA interactions.

#### 1.4.3.1 Filamentous Bacteriophages

Although different bacteriophages can be used in the phage display technique, the M13 filamentous phage is the most used (Ebrahimizadeh and Rajabibazl, 2014). Filamentous phages contain a circular single-stranded (ss) DNA genome encased in a long protein capsid cylinder. Among them, the Ff phages are the best characterized. This group contains f1, fd and M13 phages, which specifically infect *E. coli* containing the F conjugative plasmid and thus express the F pilus (Ebrahimizadeh and Rajabibazl, 2014).

Ff phages have five coat proteins. The major coat protein is pVIII, it has 50 amino acid residues and is present in almost 2700 copies in the mature phage, forming the filament tube that encapsulates the DNA. The ends of the filament are built by two different pairs of proteins; the distal end is capped by pVII and pIX, 5 copies of each. At the proximal end, pVI and pIII are present in 4 to 5 copies each (Carmen and Jermutus, 2002; Ebrahimizadeh and Rajabibazl, 2014). pIII has 406 amino acid residues and is responsible for phage infection, which is initiated by the specific attachment of pIII to the cell F pilus and ultimately results in the translocation of the phage genome into the cell cytoplasm. After replication, which involves both phage- and host-derived proteins, the viral particles are assembled in the periplasmic environment and then continuously released from the bacterial cell without lysing it. pIII is also related to the release process. Although the cell is not destroyed, its growth rate significantly decreases due to utilization of the bacterium machinery by the phage (Carmen and Jermutus, 2002; Ebrahimizadeh and Rajabibazl, 2014; Hoogenboom et al., 1998).

Ff phages reunite many other characteristics that make them appropriate for phage display applications: they are F-pilus dependent, and thus bacterial infection conditions can be controlled. Also, due to the fact that the F pilus is depolymerized after the infection, each bacterium can only be infected by one phage, resulting in the representation of one specific phage encoding a unique peptide or antibody in each bacterial clone. Finally, filamentous phages are easy to manipulate and resistant to extreme conditions such as acidic pH, high temperatures, and enzymatic cleavage, which makes them adaptable to the panning process and even suitable for *in vivo* applications (Ebrahimizadeh and Rajabibazl, 2014).

All Ff phage coat proteins have been used as fusion proteins for phage display, but the most common choices are pIII and pVIII, both coded by a single gene, respectively *gIII* or *gVIII* (Bazan et al., 2012; Carmen and Jermutus, 2002). The main differences when considering these two options are the length and the density or copy number of the displayed foreign proteins in the progeny phages. Compared to pVIII, the pIII can tolerate considerably longer peptide fusions, but given that there are only 4 or 5 copies of pIII protein per viral particle, the maximum amount of displayed fusion proteins is 5 copies per phage. If the aim is to display a higher number of peptides or small protein fragments pVIII fusion partner is recommended, once it can append up to hundreds or thousands of copies, depending on the display system used. However, the increase in the copies displayed highly restricts the length of displayed peptides toleration that could decrease to as little as 5-12 amino acid residues (Ebrahimizadeh and Rajabibazl, 2014; Scott, 2001; Qi et al., 2012). pVIII fusions multivalent display is then suitable for tests such as avidity assessment, protein–protein interactions, immunological assays or whenever enhancement of the detection signal is needed, once it results in selection of high avidity

but possibly low-affinity binders. On the other hand, type III display is most appropriate for experiments such as antibody isolation, resulting in selection of high-affinity binders in the panning processes (Ebrahimizadeh and Rajabibazl, 2014; Qi et al., 2012).

#### 1.4.3.2 Phagemid vectors

Initially, phage vectors that carried all the genetic information required for the phage life cycle were used, but phagemids have become the most widely used vector system for display (Hoogenboom et al., 1998).

Phagemids are plasmids that encode both plasmid and phage origins of replication, a phage packaging signal and an antibiotic resistance marker. Phage display appropriate phagemids also carry a cloning restriction site between sequences encoding the signal sequence and the phage coat protein gene (Rakonjac et al., 2011; Scott, 2001). Phagemids are generally smaller than phage genomes (≈4,6 kilobases), so they have higher transformation efficiencies, facilitating the construction of large and highly diverse repertoires. These vectors can accommodate larger foreign DNA fragments and usually are genetically more stable than recombinant phages under multiple propagations (Carmen and Jermutus, 2002; Qi et al., 2012; Scott, 2001). A phagemid cannot produce infective phage particles alone, once the genes essential for phage replication and assembly are missing. As such, the use of a helper phage is essential for phagemid systems since it will supply all the other proteins required to make functional phages. For that, cells already containing the phagemid vector are superinfected with a helper phage. These are normal Ff phages with a number of modifications: they usually carry antibiotic resistance genes, they contain a defective origin of replication and their packaging signal is severely disabled so that the phagemid genome is preferentially replicated and packaged (Carmen and Jermutus, 2002; Scott, 2001).

The phagemid system is one of the alternatives to overcome limitations like the lack of appropriate restriction enzyme recognition sites in the vicinity of the coat protein genes in wild-type phages. Also, this avoids the loss of coat protein functionality that could result from using the unique copy of that protein-encoding gene in the phage genome for the insertion of foreign DNA; in this case, the fusion protein will be provided by the phagemid and the wild-type version of coat protein by the helper phage (Bazan et al., 2012; Hoogenboom et al., 1998).

##### 1.4.3.2.1 pCOMB3XSS phagemid vector

pCOMB3X vectors were developed at The Barbas Laboratory, The Scripps Research Institute®, La Jolla, CA, USA. They were designed for Fab (and other proteins) display on the surface of filamentous phage, and also to allow their expression as soluble proteins (Rader and Barbas III, 1997). In those vectors the proteins are fused to the C-terminal domain of the phage coat protein pIII (Scott, 2001).

Fab library construction can be accomplished using a single *Sfi*I cloning step, given that pCOMB3X vectors include two asymmetric *Sfi*I restriction sites that allow directional cloning. Alternatively, it can be done by sequential cloning of the LC using *Sac*I and *Xba*I enzymes followed by HC Fd fragment cloning with *Xho*I and *Spe*I (Scott, 2001). The pCOMB3XSS variant includes two stuffers, a 1200 bp stuffer in the LC cloning region and a 300 bp stuffer in the HC cloning region (The

Scripps Research Institute, 2014). pCOMB3XSS holds a single *lacZ* promoter and a combination of *ompA* and *pelB* leader sequences to direct the expression of antibody LC and HC Fd-pIII fusion proteins, respectively (Scott, 2001).

Production of soluble, isolated Fab can be done by proteolysis of the labile fusion protein or by excision of the *gIII* portion of the fusion fragment from the phagemid genome using *Spe*I and *Nhe*I restriction enzymes. A third hypothesis is propagating the phagemid in a non-suppressor bacteria strain (e.g., TOP10) to turn-of the expression of the fusion protein, once this vector contains an amber stop codon (TAG) at the 5' end of *gIII*. Naturally, this implies that *gIII* fusion proteins will only be produced in suppressor strains, like those containing *supE* (e.g., TG1) or *supF* genes (Scott, 2001), in which mutant tRNA inserts a specific amino acid at UAG codons, thus suppressing its translation termination effect (Sigma-Aldrich, 2014).

pCOMB3XSS also carries an ampicillin-resistance gene and two peptide tags at the C-terminus of the displayed protein: a hexa-histidine tail ($His_6$) that facilitates purification, and the influenza hemagglutinine (HA) epitope tag that facilitates detection of the protein using an anti-HA Ab (Scott, 2001). A sketch of the vector is shown in Figure 1.7.

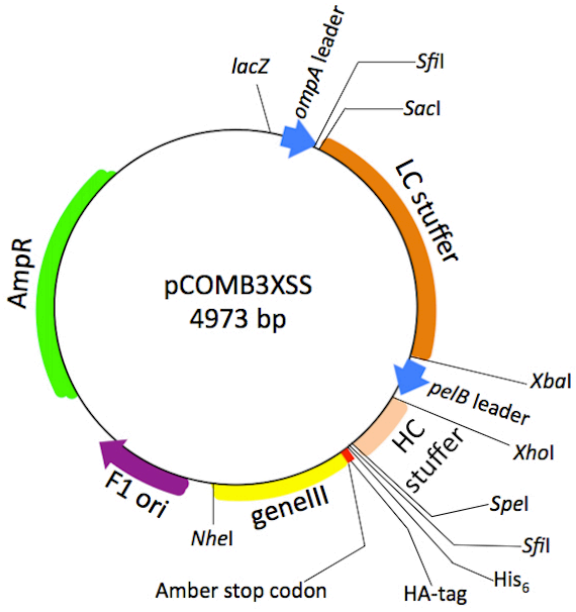


**Figure 1.7 - pCOMB3XSS phagemid vector sketch. Adapted The Scripps Research Institute, 2014.**

## 1.5 Aim of the Thesis

The aim of this thesis was the generation of an immune phage display Fab library for the development of new therapeutic Abs for breast cancer. This included the following secondary aims:

- Construction of two Fab libraries (κ and λ) by recombining the IgG repertoire of several breast cancer donors.
- Characterization of the library by assessing its size and diversity.
- Validation of the library by selection of Fab's for a general target – BSA – and a well established breast cancer target – HER2.
- Selection of specific Fab's against a target protein that intervenes in the Notch1 signaling pathway – the DLL1 ligand.

# 2. Materials and Methods

## 2.1 Materials

Restriction enzymes, T4 DNA ligase, *Taq* and *Pfu* DNA polymerase were purchased from Thermo Scientific, except *Bst*NI that was from New England Biolabs. All primers used in this work are described in Table 6.1 (appendix, section 6.1). Primers used for Ab fragments construction were designed based on previously described primers for Ab libraries construction (Andris-Widhopf et al., 2011; de Haard et al., 1999; Marks et al., 1991; Shen et al., 2007; Zhu and Dimitrov, 2009) and synthesized by STAB Vida. Milli-Q purified (MQ) water was used for molecular biology procedures. DPBS (GIBCO® Life Technologies™) was used for mononuclear cells processing and for protein, Ab and phage dilutions, unless specified otherwise. 2YT or LB medium were used for bacterial growth and dilutions. 2YT or LB agar plates with 100 µg/mL ampicillin (Sigma-Aldrich®) and 1% glucose (Sigma-Aldrich® or Applichem) (amp/glu) were used for bacteria selection in solid medium. 2YT or LB agar plates with 50 µg/mL kanamycin (Sigma-Aldrich®) and 1% glucose (Sigma-Aldrich® or Applichem) (kan/glu) were used as controls for detection of helper phage-infected TG1 cells.

### 2.1.1 Strains and plasmids

The *E. coli* strains used were: TG1 (Lucigen®) - F' [traD36 proAB+ lacIq lacZΔM15]supE thi-1 Δ(lac-proAB) Δ(mcrB-hsdSM)5, (rK-mK); TOP10 (Invitrogen™) - F- mcrA Δ(mrr-hsdRMS-mcrBC) φ80lacZΔM15 ΔlacX74 nupG recA1 araD139 Δ(ara-leu)7697 galE15 galK16 rpsL(StrR) endA1 λ-.

The plasmid used for construction of the phage libraries was pCOMB3XSS, and it was kindly provided by Carlos F. Barbas III, The Scripps Research Institute, La Jolla, CA, USA.

The helper phage used for the libraries phagemid rescue was M13KO7 (Invitrogen™).

## 2.2 Molecular Biology General Procedures

### 2.2.1 Nucleic acid quantification and purity assessment

Nucleic acid quantification and purity assessment were performed by spectrophotometry using NanoDrop ND-1000 spectrophotometer (Alfagene) or NanoVue™ (GE Healthcare).

### 2.2.2 Agarose Gel Electrophoresis

Nucleic acid samples were run in 1% agarose (Lonza) gel in 1x Tris-acetate-EDTA (TAE) buffer (5 Prime) in $dH_2O$ at 80V-120V, unless stated otherwise. Loading buffer used was TrackIt™ Cyan/Orange Loading Buffer 6x (Life Technologies™), and molecular weight marker was GeneRuler 1 kb DNA Ladder (Thermo Scientific). Gels were stained by incubating in a solution of 0,1 M NaCl (Sigma-Aldrich®) and 3x Gel Red (Biotium) with agitation. Gels were observed and photographed on GelDoc™ XR+ System (Bio Rad).

### 2.2.3 DNA Fragments Separation and Purification by Agarose Gel Electrophoresis

DNA preparation, gels run and staining were performed as described in section 2.2.2, except that lower voltages were generally used (70V-90V). Lanes containing the ladder and a small volume of the sample to be purified (typically 2-6 µL, depending on its concentration) were cut and separated from the main gel and exposed to UV light in Benchtop UV Transiluminator (UVP). This allowed the identification of the exact position of the DNA band of interest, whose expected size was known. That position was marked with a scalpel and used to guide the cut of gel bands in the same position on the lanes that contained the DNA to be purified. DNA contained in the bands was purified with illustra™ GFX PCR DNA and Gel Band Purification Kit (GE Healthcare) following the manufacturer's protocol for purification of DNA from TAE and TBE agarose gels. The following modifications were performed: agarose bands were melt with 450 rpm agitation, and DNA elution was carried out with 50 µL of MQ water. This method was applied to avoid nicking of DNA that can be caused by exposing it to the UV light and can affect subsequent manipulation steps (Clark, 2002).

### 2.2.4 DNA Purification from Solutions

DNA Purification from Solutions was performed with illustra™ GFX PCR DNA and Gel Band Purification Kit following the manufacturer's protocol for purification of DNA from solution or an enzymatic reaction, except that DNA elution was made with MQ water.

### 2.2.5 Plasmid isolation by Miniprep

Bacterial clones were grown overnight (O/N) at 37º C, 250 rpm in 3 mL-7,5 mL of medium with 100 µg/mL ampicillin, and the cultures used for phagemid isolation either with NZYMiniprep kit (NZYTech) or GeneJET Plasmid Miniprep Kit (Thermo Scientific), following the respective manufacturer's instructions, except that DNA elution was performed with MQ water.

### 2.2.6 Electrocompetent Cells – Preparation and Electroporation

An O/N culture of *E. coli* TG1 was diluted in medium to an initial $OD_{600nm}$=0,05 and incubated at 37ºC, 250 rpm, until reaching an $OD_{600nm}$≈0,5 (mid-exponential growth phase). The culture was cooled on ice for 15'-30' and centrifuged at 3220 x *g* for 20'-40' at 4º C. The pellet was resuspended in the same volume of chilled sterile MQ water and the centrifugation repeated. After this, pellet was again resuspended but in 1/2 vol of chilled sterile MQ water, and the centrifugation step repeated. The pellet was resuspended in 1/50 vol of chilled sterile 10% glycerol (Scharlau) and the centrifugation step repeated. Finally, the pellet was resuspended in 1/400 vol of chilled sterile 10% glycerol and divided in 100 µL aliquots that were kept on ice and immediately used for electroporation.

Electroporation was performed using Eporator® electroporator (Eppendorf) and 2 mm gap electroporation cuvettes (Bio Rad) with the following settings: 25 µF, 2,5 kV and 200 Ω. After the electrical impulse, the transformed cells volume was made up to 1 mL with warm medium and the cells incubated at 37 °C, 250 rpm, for 45'-1h before plating.

## 2.3 RNA Isolation

Peripheral blood samples from 16 breast cancer patients were collected to heparinised tubes at CUF Descobertas Hospital. These included 15 female adults with ages between 48 and 84 and one male with age of 71. All patients signed an informed consent form prior to donation.

For mononuclear cells isolation, the blood samples were added to the same volume of Histopaque®-1077 (Sigma-Aldrich®) solution and centrifuged at 1260 x *g* for 90' at RT with no break. Buffy coats were washed by adding 10 mL of DPBS and centrifuging at 250 x *g* for 20' at RT. Washing steps were repeated twice with 5 mL and 10 mL of DPBS, respectively. Pellets were resuspended in 500 µL of DPBS.

Mononuclear cell suspensions were centrifuged at 400 x *g* for 5'. Total RNA was extracted from pellets using GeneJET RNA Purification Kit (Thermo Scientific) following the manufacturer's protocol for Mammalian Cultured Cells Total RNA Purification, starting on step 2. Extracted RNA was treated with DNase TURBO™ DNA-free Kit (Ambion®) following the manufacturer's protocol. Total RNA was quantified and its quality assessed by agarose gel electrophoresis.

## 2.4 cDNA synthesis

A RNA pool was formed by adding the same quantity of total RNA from each donor's sample. The RNA pool was used to synthesize specific first strand cDNA using primers for IgG-derived HC CH1 region, κ LC and λ LC conserved regions mRNA. 1 µg of RNA was used to produce HC cDNA in 120 µL of reaction with 20 pmol of primer IgG CH1 Rv. Each LC cDNA was synthesized from 0,76 µg of RNA in 75 µL of reaction. For κ cDNA, 20 pmol of primer CLκ Rv were used, and for λ cDNA, 10 pmol of primer CLλa Rv plus 10 pmol of primer CLλb Rv. Additionally, each reaction contained 200 U of RevertAid Reverse Transcriptase (Thermo Scientific), 1x reaction buffer, 20 U of Ribolock RNase Inhibitor (Thermo Scientific) and 1mM of dNTP mix (NZYTech). Reactions were incubated at 42º C for 60'. The enzyme was inactivated by heat, incubating at 70º C for 10'.

## 2.5 Light Chain and Heavy Chain Fd Fragments Construction

PCRs were performed in My Cycler™ Thermal Cycler (Bio Rad) in a volume of 50 µL containing the following components: 10 ng of template (except for cDNA template; in reactions with 2 different templates, 5 ng of each were used), 0,2 mM of dNTPs, 20 pmol of each primer (forward, Fw, and reverse, Rv), 1,25 U of *Pfu* DNA polymerase (Thermo Scientific) and 1x buffer. Different annealing temperatures were set depending on the primers used in each reaction. Extension time varied accordingly to each product expected size, being at least 2'/kb, which is the recommended for *Pfu* DNA polymerase. Other conditions were common, namely: initial denaturation at 95º C for 2', followed by 30 cycles of denaturation at 95 °C for 30'', annealing for 30'', and extension at 72º C; and 1 last cycle of final extension at 72º C for 10'. A "no template" control was performed for all reactions. 5 µL of each PCR product was analyzed by electrophoresis on 1% agarose gel. Subsequently, the remaining volume of all successfully amplified products (generally 45 µL) was gel purified before being further manipulated. All exceptions to those conditions are indicated.

### 2.5.1 Amplification of VH gene repertoire

VH gene repertoire was constructed in three essential steps, represented in Figure 2.1. In the first step, HC cDNA was used as template for VHCH1 amplification. Two sets of VHCH1 fragments were prepared:

A) 4,8 µL of template were used per reaction. 6 separate reactions were performed using 6 different Fw primers (VH1-6 Fw) and the same Rv primer (IgG CH1 Rv). Reactions were prepared with 10 pmol of each primer. Annealing temperature was 59º C and extension time 1,5'. Exception was VH6CH1 product, which was amplified in two separate reactions with different annealing temperatures, namely 53,9º C and 52,5º C.

B) 1,2 µL of template were used per reaction. The same primer combinations of set A were used, but reactions contained 20 pmol of each. Annealing temperature was 59º C and extension time 2'.

In the second step, the first step products were used as template for CH1 (set A) and VHJH (set B) segments amplification. A different reaction was prepared with each type of template. For CH1 amplification, 10 pmol of each primer and 40 ng of template were used per reaction, except for VH4CH1 (32 ng) and VH5CH1 (33 ng). Fw primer was CH1 Fw and Rv primer IgG CH1 Rv. Annealing temperature was 55º C and extension time 26''. For VHJH amplification, 2,5 µL of template and 20 pmol of each primer were used per reaction. Fw primer was the same used for the respective VHCH1 template production. 4 Rv primers were used (JH 1/2, 3, 4/5 and 6 Rv) for each different template, so 24 different reactions were performed. Annealing temperature was 57º C and extension time 1,5'.

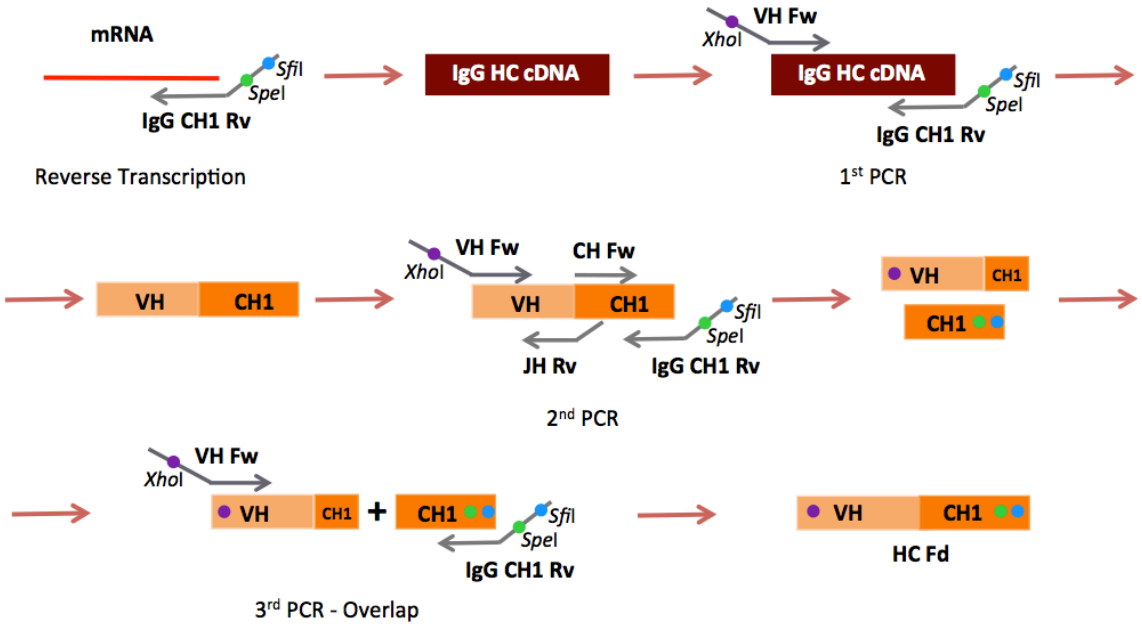


**Figure 2.1 - PCR amplification steps for VHCH1 (HC Fd) fragments construction. Adapted from a scheme provided by Inês Barbosa.**

The third step consisted on reamplification and assembly of the produced VHJH and CH1 fragments by overlap PCR. A CH1 pool was used as CH1 template and was formed by adding the same quantity of all 6 CH1 products from step 2. 6 different pools were formed with the VHJH products. For each, the same quantity of the 4 products obtained from the same template was added.

Each VHJH pool was combined with the CH1 pool in a different reaction. Fw primer was the same used for the production of the fragments that composed the respective VHJH pool and Rv primer was IgG CH1 Rv. Two overlap PCRs were performed with different combinations of annealing temperature and extension time, namely 55º C and 2', or 59º C and 1,75'. The products (VHCH1 – Fd – fragments) obtained from the same template by both amplifications were added before gel purification.

The 6 VHCH1 (Fd) fragments were reamplified from the respective VHCH1 (Fd) purified overlap products, using the same primers used for the overlapping PCR. Annealing temperature was 59º C and extension time 1,75'.

### 2.5.2 Amplification of κ LC gene repertoire

κ LC repertoire was constructed in two steps, represented in Figure 2.2. In the first, κ LC cDNA was used as template to separately amplify the VLκ and CLκ domains, namely 3 µL per reaction. For VLκ amplification, 9 separate reactions were performed using 9 different Fw primers (VLκ1, 2, 3a, 3b, 3c, 3d, 4, 5 and 6 Fw) and the same Rv primer (VLκ Rv). CLκ region was amplified with primers CLκ Fw and CLκ Rv. Annealing temperature was 50º C and extension time 1'. Exception was VLκ5 product, which was amplified in two separate reactions with different annealing temperatures, namely 48,1º C and 47,1º C.

The second step consisted on reamplification and assembly of the produced VLκ and CLκ fragments by overlap PCR. Each VLκ product was combined with the CLκ fragment in a different reaction. Fw primer was the same used for the respective VLκ fragment production and Rv primer was CLκ Rv. Annealing temperature was 50º C, except in the reactions with VLκ4 Fw and VLκ5 Fw primers, which was 48º C. Extension time was 1,67'.

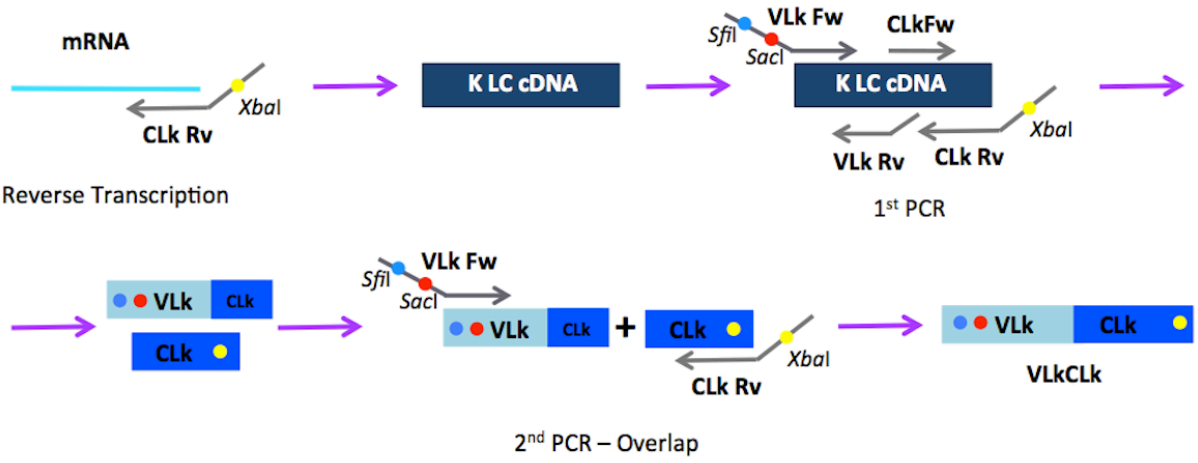


**Figure 2.2 - PCR amplification steps for VLκCLκ fragments construction. Adapted from a scheme provided by Inês Barbosa.**

### 2.5.3 Amplification of λ LC gene repertoire

λ LC repertoire was constructed in two steps (Figure 2.3). In the first, λ LC cDNA was used as template to separately amplify the VLλ and CLλ domains, namely 3 µL per reaction. For VLλ amplification, 12 separate reactions were performed using 12 different Fw primers (VLλ1a, 1b, 1c, 2, 3a, 3b, 4, 6, 7/8a, 7/8b, 9 and 10 Fw) and the same Rv primer (VLλ Rv). For CLλ region amplification,

2 separate reactions were performed using the same Fw primer (CLλ Fw) and 2 different Rv primers (CLλa Rv and CLλb Rv). Annealing temperature was 62º C for VLλ fragments amplification and 50º C for CLλ's. Extension time was 1'. Exception was VLλ10 product, which was amplified in three separate reactions with different annealing temperatures, namely 55º C, 53º C and 50,9º C.

The second step consisted on reamplification and assembly of the produced VLλ and CLλ fragments by overlap PCR. Each VLλ product was combined with each CLλ fragment in a different reaction, making a total of 24 reactions. Fw primer was the same used for the respective VLλ fragment production and Rv primer the one used for the respective CLλ fragment production. Annealing temperature was 50º C. Extension time was 1,5'.

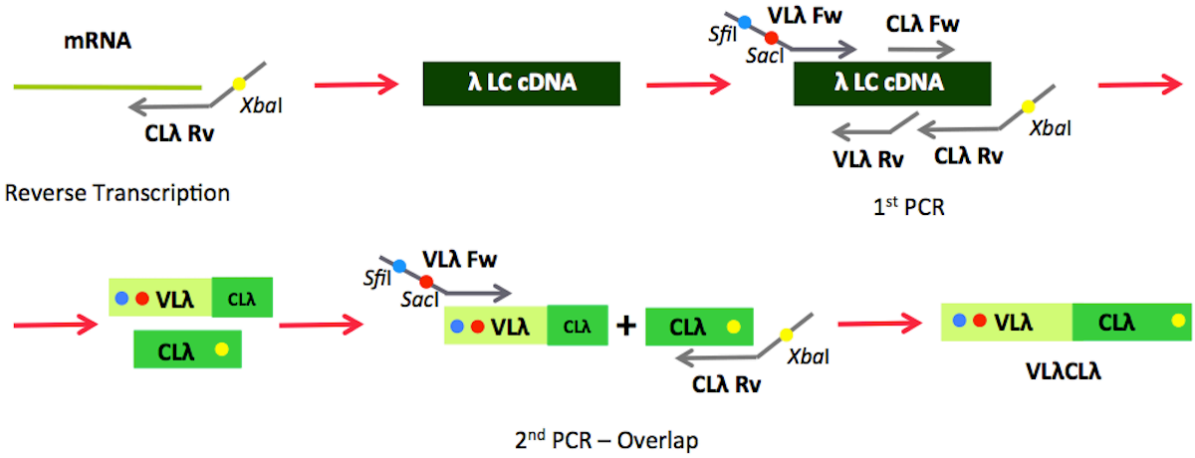


**Figure 2.3 - PCR amplification steps for VLλCLλ fragments construction. Adapted from a scheme provided by Inês Barbosa.**

## 2.6 Libraries Construction

Restriction enzyme digestions were performed following the manufacturer's instructions. Briefly, reactions contained 0,3 µg-1,5 µg of DNA, 1x recommended buffer and excess of enzyme (at least 10 U per 1 µg of DNA). Final volume varied between 20 µL-100 µL, once upscale was performed when necessary. Reactions were performed at 37º C for 1h-4h.

Dialysis was performed with MF$^{TM}$ membrane filter (Millipore) 0,025 µm VSWP. Products to be dialysed were deposited on a membrane floating on distilled water in a petri dish placed on ice, for 1h.

Prior to ligation, digested vectors were gel purified as described in section 2.2.3, quantified and dephosphorylated. Dephosphorylation was performed with FastAP Thermosensitive Alkaline Phosphatase (Thermo Scientific) following the manufacturer's instructions but for longer times (1h-3h). Briefly, reactions contained 0,2 µg-4 µg of DNA, 1x buffer and excess of enzyme (1 U-2 U per 1 µg of DNA). Final volume varied between 20 µL-80 µL, once upscale was performed when necessary.

Digested PCR products were purified and quantified prior to ligation, as well as ligation products before large scale transformation, as indicated in section 2.2.1. Ligation reactions were performed with T4 DNA ligase (Thermo Scientific) following the manufacturer's instructions. Briefly, reactions contained 0,11 µg-3 µg of DNA, 1x buffer and excess of enzyme (at least 5 U per 1 µg of DNA). Final volume varied between 20 µL-300 µL, once upscale was performed when necessary.

Reactions were performed at RT for 1,5h-5h at RT, and an excess of enzyme was used. Small scale trial ligations were performed with different verctor:insert molar ratio (calculated following Formula 2.1) and as well as ligation reactions containing dephosphorylated vector but no insert were also performed.

$$\frac{vector\ quantity\ (ng) \times insert\ size\ (kb)}{vector\ size\ (kb)} \times molar\ ratio \frac{insert}{vector} = insert\ quantity\ (ng)$$

**Formula 2.1 – Insert quantity calculation considering the verctor:insert molar ratio.**

Ligations were transformed by electroporation into electrocompetent *E. coli* TG1 cells freshly prepared as described in section 2.2.6. Every time, 10 pg of the plasmid pUC19 (Lucigen) was also electroporated into the bacterial cells in order to calculate the transformation efficiency, and for vector self-ligation transformation, 10 µL of digested vector was used. Aliquots from each transformation were plated in amp/glu plates. 100 µL of untransformed cells were plated in amp/glu and kan/glu plates for control purposes. Plates were incubated O/N at 37 °C. For small scale trials, colonies were counted and the number of transformants per µg of DNA calculated following Formula 2.2. This allowed determining the optimal verctor:insert molar ratio to create the libraries and to assess the background levels of vector self-ligation.

$$\frac{\#colonies \times total\ transformed\ cells\ volume\ (\mu L) \times 10^6}{transformed\ DNA\ quantity\ (pg) \times plated\ volume\ (\mu L)} = \#transformants/(\mu g)\ DNA$$

**Formula 2.2 – Number of transformants (#transformants) per µg of transformed DNA calculation.**

For the libraries construction (large scale), ligation reactions were prepared with a 1:5 vector:insert molar ratio. Several transformations of 1 µL-2 µL of DNA into 100 µL of cells were performed for each ligation. The number of colonies in the test plates was used to calculate the library size. The remaining transformation volumes were pooled together and processed as specified ahead.

### 2.6.1 Cloning of the LC repertoires

For LC repertoires cloning, *Sac*I and *Xba*I enzymes (Thermo Scientific) were used. Prior to digestion, a VLκCLκ pool was formed by adding the same quantity of all 9 VLκCLκ PCR products. For the λ repertoire, two pools were prepared, namely for VLλCLλa and VLλCLλb fragments, each containing the same amount of the respective 12 PCR products. For VLκCLκ pool a double digestion was performed. For VLλCLλ pools and pCOMB3XSS vector, sequential digestions were performed. *Sac*I digestion products were dialysed before digestion with *Xba*I. After purification and quantification, the same quantity of digested VLλCLλa and VLλCLλb were pooled together.

VLλCLλ**Sac*I/*Xba*I pool and VLκCLκ**Sac*I/*Xba*I were ligated into pCOMB3XSS**Sac*I/*Xba*I in separate reactions. For VLκCLκ sub-library construction, 23 transformations of 2 µl of pCOMB3XSS+VLκCLκ ligation were performed. For VLλCLλ sub-library construction, 23 transformations of 1 µl of pCOMB3XSS+VLλCLλ ligation were performed.

For each sub-library, 5 mL of the transformation pool were plated in 10 large amp/glu plates (Nunclon$^{TM}$Δ 245x245x25 mm, Nunc) (500 µL/plate) and incubated O/N at 37º C. Colonies on the

large plates were scraped off using medium with 25% glycerol. Resulting pool was aliquoted and then stored at -80º C as bacterial stock.

The remaining pool volume of the transformation was made up to 100 mL with medium and incubated at 37 °C, 250 rpm, for 1h. Ampicillin was added to a final concentration of 100 µg/mL and the culture incubated at 37 °C, 250 rpm, for 1h. 400 mL of medium with ampicillin at 100 µg/mL were added to the culture and it incubated O/N at 37 °C, 250 rpm. Suspension culture was centrifuged for 30' at 3220 x *g* and the pellets used to prepare phagemid DNA with NZYMaxiprep kit (NZYTech) following the manufacturer's instructions for low copy number plasmids isolation starting on step 2. DNA pellets were reconstituted in MQ water.

### 2.6.2 Cloning of the VH repertoire

For VH repertoire cloning *Xho*I and *Spe*I enzymes (Thermo Scientific) were used. Prior to digestion, a VHCH pool was formed by adding the same quantity of all 6 VHCH PCR products. VHCH pool, pCOMB3XSS+VLκCLκ and pCOMB3XSS+VLλCLλ vectors were digested with *Xho*I, dialysed and then digested with *Spe*I.

VHCH***Xho*I/*Spe*I was ligated into pCOMB3XSS+κ***Xho*I/*Spe*I and pCOMB3XSS+λ***Xho*I/*Spe*I in separate reactions. For VLκCLκ-VHCH library construction (κ library), 19 transformations of 2 µl of pCOMB3XSS+VLκCLκ+VHCH ligation were performed. For VLλCLλ-VHCH library construction (λ library), 17 transformations of 2 µl of pCOMB3XSS+VLλCLλ+VHCH ligation were performed.

For each library, the transformation pool was centrifuged for 30' at 3220 x *g*. Resuspended culture was approximately 3 mL; this volume was plated in 6 large amp/glu plates (~500 µL/plate and incubated O/N at 37º C. Colonies were scraped off into medium with 25% glycerol. Resulting pool was aliquoted and then stored at -80º C as bacterial glycerol stock.

### 2.6.3 Libraries Phage Rescue

Phage rescue protocol was adapted from de Haard, 2002. Aliquots of both libraries bacterial glycerol stock were separately diluted in 400 mL of medium with 1% glucose (Sigma-Aldrich®) and 100 µg/mL ampicillin to an initial $OD_{600nm}$=0,05. Cultures were incubated at 37ºC, 200 rpm, until reaching an $OD_{600nm}$≈0,5 (growth phase in which F piliation reaches a maximum (Kotlan and Glassy, 2009)). At this stage 1,6 x $10^{12}$ cfu of helper phage was added to each culture, so that the multiplicity of infection (MOI) would be 20:1 (#phage particles:#host cells) (Marks and Bradbury, 2004a). The relation between the $OD_{600nm}$ and the number of viable TG1 bacteria was determined following the growth curves. During the infection step with the helper phage cultures were incubated statically at 37º C for 40', plus 50' with shaking at 100 rpm (de Haard, 2002). 320 mL of each culture were centrifuged at 3220 x *g* for 25'. Pellets were resuspended in the same volume of medium with 100 µg/mL ampicillin and 25 µg/mL kanamycin, and incubated O/N at 30º C, 250 rpm. On the following day, cultures were centrifuged at 3220 x *g* for 30'. 64 mL of 20% PEG 2,5 M NaCl (1/5 vol) was added to each culture's supernatant, and phages precipitated on ice for 1h. Phage suspensions were centrifuged at 6200 x *g* for 20' at 4º C. A second precipitation step was carried out, in which phage pellets were resuspended in 32 mL of DPBS (1/10 vol) and 6,4 mL of 20% PEG 2,5 M NaCl (1/50 vol) was added. Phages precipitated on ice for 1h and centrifugation was repeated. Pellets were resuspended in 15 mL of

DPBS and filtered with a 0,45 µm filter (Millipore) to remove any residual bacteria. In the final precipitation step 3 mL of 20% PEG 2,5 M NaCl (1/5 vol) was added and phages precipitated on ice for 1h. Centrifugation was repeated and final pellets were resuspended in 3,2 mL of DPBS with 10% glycerol (1/100 vol of the initially centrifuged culture). Phages were aliquoted and stored at -80º C.

One aliquot from each library was used to titer and determine the phage libraries size. An O/N grown *E. coli* TG1 culture was diluted in medium to an initial $OD_{600nm}$=0,05 and incubated at 37ºC, 250 rpm, until reaching an $OD_{600nm}$≈0,5. Serial dilutions of each phage library were prepared in medium. 90 µL of TG1 culture at $OD_{600nm}$≈0,5 was added to 90 µL of each phage dilution and incubated at 37º C for 30', 50 rpm. 10 µL of dilutions $10^{-5}$ to $10^{-12}$ were plated in amp/glu plates. Uninfected TG1 cells were also plated in amp/glu and kan/glu plates for control purposes. Plates were incubated O/N at 37º C. Colonies were counted and cfu/mL were calculated following Formula 2.3.

$$\frac{\#colonies \times dilution\ factor}{plated\ volume\ (\mu L)} \times 1000 = cfu/mL$$

**Formula 2.3 – Colony forming units (cfu) per mL calculation from colonies count on growth plates.**

## 2.7 Libraries Characterization and Quality Assessment

### 2.7.1 Colony PCR

Serial dilutions of Fab κ and λ libraries bacterial glycerol stocks were prepared in medium, spread in amp/glu plates and incubated O/N at 30º C. The inserts presence was screened by PCR on 100 randomly selected isolated colonies of each library. A pinch of each colony was directly picked from the agar plate and added to a reaction mixture containing 0,4 mM of dNTPs, 20 pmol of pCOMB Fw primer, 20 pmol of pCOMB Rv primer, 2 mM $MgCl_2$ and 1x buffer. Mixtures incubated at 95º C for 10' and 1,25 U of *Taq* DNA polymerase (Thermo Scientific) was added to the reaction, with a reaction final volume of 50 µL. PCR was performed with the following conditions: initial denaturation at 95º C for 2', followed by 30 cycles of denaturation at 95 °C for 30'', annealing at 60º C for 30'', and extension at 72º C for 2'; and 1 last cycle of final extension at 72º C for 10'. The following controls were performed: "no template" reaction; untransformed TG1, pCOMB3XSS, pCOMB3XSS+VLκCLκ and pCOMB3XSS+VLλCLλ libraries' DNA amplification. 5 µL of each PCR product was analyzed by DNA gel electrophoresis as described in section 2.2.2.

### 2.7.2 DNA Fingerprinting

The DNA fingerprinting profile of 44 κ and 47 λ clones that presented the correct size fragment of the constructed Fab by the PCR was analyzed. For that, 5 µL of these selected PCR products were digested with 1 µL of *Bst*NI restriction enzyme (New England Biolabs) in a 20 µL reaction for approximately 2h at 60º C. 9 µL of each digestion product was analyzed by electrophoresis as described in section 2.2.2. As an exception the agarose gel concentration was 3% to allow separation of the very small fragments generated from the *Bst*NI digestion, and the run was performed at 80V for about 2h.

### 2.7.3 DNA Sequencing

Clones whose DNA fingerprinting profile was analyzed were also sequenced. Clones' phagemid DNA was extracted by miniprep as described in section 2.2.5. DNA was quantified as described in section 2.2.1 and several samples were run in 0,8%-0,9% agarose gel electrophoresis (see section 2.2.2). Sequencing was performed at STAB Vida or GATC Biotech with pCOMB Fw and pCOMB Rv primers. Sequences were submitted to VBASE2 database analysis tool (Retter et al., 2005) for sequence analysis and V-genes identification.

### 2.7.4 ELISA

Enzyme-linked immunosorbent assays (ELISA's) were executed in 96-well Nunc Maxisorp plates. Plates were coated with 50 µL of protein or Ab at 5 µg/mL and incubated O/N at 4º C or for 1h-1h20' at room temperature (RT). After coating plates were washed 3 times with 300 µL of PBST 1x in WellWash Versa (Thermo Scientific). Wells were blocked with 250 µL of PBST-MP 3% for 1h-3h20' at RT. 50 µL of testing solution was added to each well and incubated for 1h-1h45' at RT. Washing step was repeated. 50 µL of detection Ab (Fab-on-phage ELISA: HRP-Anti-M13 Ab, 1:5000 – GE Healthcare; sFab ELISA: HRP-Anti-Human IgG Ab, 2,5 µg/mL – Sigma-Aldrich®) were added to the wells and incubated for 1h at RT. Washing step was repeated. 50 µL of RT colorimetric substrate for horseradish peroxidase 3,3',5,5'-tetramethylbenzidine (TMB, Invitrogen™) was added to the wells and incubated in the dark before the color reaction became saturated, or until negative controls started to react. Reaction was stopped by addition of 50 µL of 2N $H_2SO_4$ (AVS Titrinorm®) per well. Plates absorbance was measured at 450 nm in Multiskan™ FC (Thermo Scientific). Exceptions to the above statements will be mentioned.

#### 2.7.4.1 Titration Fab-on-Phage ELISA

Two assays were performed. In each, 1:2 serial dilutions of each library, starting at $1x10^{12}$ cfu/mL to $7,81x10^{9}$ cfu/mL, were tested in duplicate. In assay 1, the coatings were: 1) anti-M13 Ab (Acris); 2) anti-Fab Ab (Sigma); 3) uncoated. In assay 2, the coatings were: 1) anti-M13, 2,5 µg/mL Ab; 2) anti-Fab Ab; 3) anti-LC Ab (anti-κ or anti-λ Ab, respectively – Sigma-Aldrich®); 4) anti-histidine Ab (GE Healthcare); 5) uncoated.

## 2.8 Libraries' Selection by Panning

### 2.8.1 Panning

Panning was performed against three targets: bovine serum albumin (BSA) for library construction validation, human HER2 for breast cancer biomarker validation, and human DLL1 (Table 2.1). Selection with HER2 and DLL1 was performed only for λ library. Three rounds were performed for each panning. The panning protocol was adapted from Rader et al., 2001.

The panning was performed on immobilized antigen. The protocol was executed in 96-well Nunc Maxisorp plates. Coatings were performed O/N at 4º C or for 1h at RT as follows: BSA – 20 µg/mL, 100 µL per well; HER2 and DLL1 – 15 µg/mL, 50 µL per well. As a depletion step, uncoated wells were used to eliminate the unspecific plastic phage binders. Coated and depletion wells

(uncoated) were blocked with 250 µL of DPBST-MP 3% for 1h-2h at RT. 50 µL or 100 µL of phages (the same volume of coating solution) were added to the depletion wells and incubated 30' at RT with gentle shaking. In the 1st round of selection, $1x10^{12}$ phages from the phage libraries glycerol stock were used. In the following rounds, 50 µL of amplified phages from the previous round were used (for BSA panning, those were diluted in 50 µL of DPBS, so that the final volume was 100 µL). Since the HER2 molecule used is Fc tagged the panning strategy for this antigen had an extra depletion step, after the general one, in a well coated with human IgG1 (Sigma-Aldrich®) at 15 µg/mL to eliminate any Fc binders. Before proceeding to selection, 20 µL – for BSA panning – or 12 µL of phages – for HER2 and DLL1 pannings – were set aside for later determination of the number of input phages. Phages were transferred from the depletion wells to the selection wells and incubated 1h at RT with gentle shaking – selection step. In the 1st round, phages were selected twice, by transferring the phages to another selection well after the first selection and repeating the process. Selection well(s) were washed with 250 µL of 0,05% PBST buffer, by pipetting up and down and incubating 5'. This process was repeated 5 times in the 1st round and 10 times in the following. Specifically bound phages were eluted by adding 50 µL or 100 µL (the same volume of coating solution) of triethanolamine (TEA, Sigma-Aldrich®) 100 mM to the selection well(s) and incubating 10' at RT with gentle shaking. After pipetting up and down, eluted phages were transferred to a tube containing the same volume of 1 M Tris-HCl buffer (GIBCO® Life Technologies™), pH 7,5 for neutralization.

**Table 2.1 - Antigens used for panning**

| Target | Supplier | Molecular Weight | Characteristics |
|---|---|---|---|
| **BSA** | Sigma-Aldrich® | 66,43 kDa | --- |
| **HER2** | Sino Biological | 96,1 kDa | Extracellular domain of human ErbB2 (Met 1 - Thr 652) fused with the Fc region of human IgG1 at the C-terminus |
| **DLL1** | Produced in house in HEK293-EBNA6 cells | 33,52 kDa | Extracellular domain portion of human DLL1 (Trp 159 – Val 444, domains DSL to EGF-like 6) fused with a polyhistidine-tag ($His_8$) at the N-terminus |

### 2.8.2 Phage Amplification

A previously grown *E. coli* TG1 culture was diluted in medium to an initial $OD_{600nm}$=0,05 and incubated at 37ºC, 250 rpm, until reaching an $OD_{600nm}$≈0,5. 125 µL – for BSA panning – or 62,5 µL – for HER2 and DLL1 pannings – of selected phages were added to 5 mL of TG1 culture at $OD_{600nm}$=0,5 and incubated 15' at RT with gently shaking for infection. 20 µL of infected bacteria were set aside for later determination of the number of output phages. 3 mL of pre-warmed medium with ampicillin (to a final concentration of 100 µg/mL) was added to the infected bacteria and incubated 1h at 37ºC, 250 rpm. Helper phage at a MOI of 20:1 and 92 mL of pre-warmed medium with ampicillin 100 µg/mL was added and incubated at 37ºC, 250 rpm. After 1,5h-2h, kanamycin was added to a final concentration of 50 µg/mL (Lee et al., 2007) and cultures incubated O/N at 37ºC, 250 rpm. On the next day, cultures were centrifuged at 3500 x *g* for 15'-20'. Supernatants were filtered with a 0,45 µm filter to remove any residual bacteria. 20 mL of 20% PEG 2,5 M NaCl (0,2 volumes) was added to each culture supernatant and phages precipitated on ice for 1h-2h. Phage suspensions were centrifuged at 6200 x

*g* for 20' at 4ºC and pellets resuspend in 2 mL of DPBS. Phages were used for the next panning rounds and/or screening.

### 2.8.3 Titration of input and output phages

For titration of input phages, the phages set aside after selection were serial diluted in medium. 90 µL of TG1 culture at $OD_{600nm}$≈0,5 was added to 90 µL of each phage dilution and incubated 15' at RT with gentle shaking for infection. Several dilutions ($10^{-8}$ to $10^{-11}$ for the 1st round and $10^{-6}$ to $10^{-9}$ for the following) were plated in amp/glu plates. For titration of output phages, bacteria set aside after infection with selected phages was serial diluted in medium. Dilutions $10^{-1}$ to $10^{-4}$ were plated in amp/glu plates. Uninfected TG1 cells were also plated in amp/glu and kan/glu plates as controls. Plates were incubated O/N at 37º C. Colonies were counted and cfu/mL were calculated following Formula 2.3 and then used to calculate the final phage input, considering the volume that was placed in the selection well (after setting apart the volume for input titration) and the phage output (considering the output infection volume, namely 5 mL). The percentage of phage recovery was calculated following Formula 2.4 and the enrichment between rounds following Formula 2.5.

$$\frac{output\ phages\ in\ round\ n}{input\ phages\ in\ round\ n} \times 100 = Phage\ recovery\ in\ round\ n\ (\%)$$

**Formula 2.4 – Percentage of phage recovery in a determined round of selection calculation, following Arbabi-Ghahroudi et al, 2009.**

$$\frac{Specific\ phages\ recovery\ in\ round\ n}{Specific\ phages\ recovery\ in\ round\ n-1} = Enrichment\ in\ round\ n$$

**Formula 2.5 – Enrichment: increase in recovered phage in a determined round of selection relatively to the previous round calculation, following Arbabi-Ghahroudi et al, 2009.**

## 2.9 Analysis of Selected Phage Pools

### 2.9.1 Phage Pools Screening by Fab-on-Phage ELISA

For pools screening by ELISA, 1:2 serial dilutions of phage pools were tested in duplicate, unless stated otherwise.

#### 2.9.1.1 BSA Selected Phages

Two assays were performed. In assay 1, there were no duplicates. Both eluted and amplified phages from round 1 of both libraries and round 2 of λ library were assayed. In assay 2, amplified phages from round 3 of the λ library and from round 2 of κ library were assayed. For both assays, coatings were: 1) anti-Fab Ab; 2) anti-LC Ab (anti-κ or anti-λ Ab, respectively); 3) BSA; 4) uncoated.

#### 2.9.1.2 HER2 Selected Phages

Two assays were performed. In assay 1, round 2 amplified phages were tested and coatings were: 1) anti-Fab Ab; 2) HER2; 3) uncoated. In assay 2, round 2 and 3 amplified phages were tested. Round 2 was tested for the following coatings: 1) HER2; 2) IgG1. Round 3 was tested for the following coatings: 1) anti-Fab Ab; 2) HER2; 3) IgG1; 4) uncoated.

#### 2.9.1.3 DLL1 Selected Phages

Two assays were performed. In assay 1, round 2 and 3 amplified phages were tested and coatings were: 1) anti-Fab Ab; 2) DLL1; 3) uncoated. In assay 2, round 3 amplified phages test was repeated for the same coatings, but more dilutions were tested.

## 2.10. Analysis of Selected Clones

### 2.10.1 Fab-on-Phage ELISA of Selected Clones

#### 2.10.1.1 Clones Rescue

Individual colonies were picked from round 3 output plates and used for phage rescue. Protocol was adapted from Steinberger, 2001. Although both libraries were selected against BSA, only clones originating from the λ library panning were rescued and individually screened.

10 λ clones selected for BSA and 20 for HER2 or DLL1 were picked. Each clone was inoculated in 5 mL (for BSA-selected clones) or 7,5 mL (for HER2- and DLL1-selected clones) of medium with 100 µg/mL of ampicillin and incubated for 4h40'-5h20' at 37º C, 250 rpm. Helper phage was added at a MOI of 20:1 to each culture and incubated for 2h at 37º C, 250 rpm. Kanamycin was added to a final concentration of 50 µg/mL and cultures incubated O/N at 37º C, 250 rpm. On the following day, cultures were centrifuged at 2800 x *g* for 15'. 20% PEG 2,5 M NaCl was added to a final concentration of 20% and phages precipitated on ice for 1h10'-1h30'. Phage suspensions were centrifuged for 20'-30'. Phage pellets were resuspended in 200 µL (for BSA-selected clones) or 350 µL (for HER2- and DLL1-selected clones) of DPBS.

#### 2.10.1.2 Clones Fab-on-Phage ELISA

Selected individual clones were tested by Fab-on-phage ELISA, in which phages were used directly from rescue. For BSA-selected clones test, coatings were: 1) anti-Fab Ab; 2) anti-λ Ab; 3) BSA; 4) uncoated. For HER2-selected clones test, coatings were: 1) anti-Fab Ab; 2) HER2; 3) IgG1; 4) uncoated. For DLL1-selected clones test, coatings were: 1) anti-Fab Ab; 2) DLL1; 3) uncoated. Duplicate tests were performed for the phages affinity against the antigen for which they had been selected.

### 2.10.2 sFab ELISA of Selected Clones

#### 2.10.2.1 Clones sFab Induction

Clones that were considered positive for specific binding on Fab-on-phage ELISA were induced to express soluble Fab's (sFab's). Each clone phagemid DNA was extracted using GeneJET Plasmid Miniprep Kit as described in section 2.2.5. 1 µL of DNA was transformed by electroporation into 50 µL of electrocompetent TOP10 cells (from glycerol stock at -80º C) as described in section 2.2.6. Transformed cells were plated in amp/glu plates, which was incubated at 37º C O/N. Following steps (Fab expression induction protocol) were adapted from Steinberger, 2001. For each clone, a streak of TOP10 transformed bacteria was inoculated in 5 mL of medium with 100 µg/mL of ampicillin and incubated for 6h30'-8h at 37º C, 200-250 rpm. IPTG (NZYTech) was added to a final

concentration of 2 mM and cultures incubated O/N at 37º C, 250 rpm. Induced cultures were centrifuged for 10' at 2000 x *g* and pellets were resuspended in 750 µL of DPBS. Cells were lysed by 3 cycles of freezing in dry ice and ethanol bath for 3'-5' and thawing at 37º C in water bath for 3'. Lysates were centrifuged at 15000 x *g* for 5' and supernatants containing Fab's were kept for further usage.

#### 2.10.2.2 Clones sFab ELISA

Clones that were considered positive for specific binding in Fab-on-phage ELISA were further tested by sFab ELISA. Supernatants obtained from centrifugation of induced bacteria lysates were directly tested in duplicates. Coatings were the same as for Fab-on-phage ELISA, but only uncoated wells were used as negative controls for HER2-selected clones and an extra coating was used for DLL1-selected clones test, namely a histidine-tagged protein to eliminate any his tagged specific binders.

### 2.10.3 Positive Clones Analysis

#### 2.10.3.1 Genetic tests

Clones' phagemid DNA was used as template for PCR. Reactions containing 20 ng of template, 0,4 mM of dNTPs, 20 pmol of pCOMB Fw primer, 20 pmol of pCOMB Rv primer, 1,25 U of *Pfu* DNA polymerase and 1x buffer were prepared in a final volume of 50 µL. PCR was performed with the following conditions: initial denaturation at 95º C for 2', followed by 30 cycles of denaturation at 95 °C for 30'', annealing at 60º C for 30'', and extension at 72º C for 3'; and 1 last cycle of final extension at 72º C for 10'. For control purposes, pCOMB3XSS amplification and a "no template" reaction were also performed.

Clones' phagemid DNA was double digested with FastDigest *Sac*I and FastDigest *Spe*I (ThermoScientific) restriction enzymes following the manufacturer's instructions for 32'.

Clones' phagemid DNA (150 ng), PCR products (5 µL) and double-digestion products (150 ng) were analyzed by electrophoresis in 0,8% agarose gel.

BSA- and HER2-selected clones' PCR products were purified with NZYGelpure kit (NZYTech) following the manufacturers instructions for PCR clean-up. Purified PCR products and phagemid DNA of all selected clones was sequenced and results analyzed as decribed in section 2.7.3. Sequences were aligned to allow identification of unique clones using ApE – A Plasmid Editor software.

DLL1-selected clones PCR products were digested with *Bst*NI enzyme as described in section 2.7.2. *Bst*NI products (9 µL) were analyzed by electrophoresis on 3% agarose gel.

#### 2.10.3.2 SDS-PAGE and Western Blot

DLL1-selected clones sFab's were run on SDS-PAGE under reducing conditions. 5 µL of supernatants obtained from centrifugation of induced bacteria lysates were added to 5 µL of XT sample loading buffer 4x (Bio Rad) and 10 µL of DPBS and incubated 5' at 99º C before being loaded in a Criterion™ XT Precast Gel (Bio Rad). NZYColour Protein Marker II (NZYTech) was used as molecular weight marker. Gel ran at 120 V for 2h in 1x XT MOPS running buffer (Bio Rab) diluted in distilled water.

Proteins were transferred to a nitrocellulose membrane using the iBlot™ Dry Blotting System (Invitrogen™) following the manufacturer's instructions to use the device with iBlot® Gel Transfer Stacks, Nitrocellulose (Invitrogen™). Blotting parameters were 20V for 7' (device P3 default program). Membrane was blocked with PBST-MP 3% for 30' with shaking and washed 3 times by shaking in PBST for 5'. Membrane incubated with HRP-anti-Fab Ab at 2,5 µg/mL in DPBS for 55' and washing was repeated. Membrane was photographed on ChemiDoc™ XRS+ System (Bio Rad) under EPI white light. Western blot revelation was performed by covering the membrane with 1 mL of Western Lightning® Plus-ECl, Enhanced Chemiluminescence Substrate (Perkin Elmer), previously prepared as indicated by manufacturer's instructions. Membrane was exposed to chemiluminescent light and photographed after different exposition times on ChemiDoc™.

## 2.11 Antigens Titration ELISA

A 96-well Nunc Maxisorp plate was coated overnight at 4º C with 1:2 serial dilutions of each antigen used for panning, ranging from 20 µg/mL to 0,16 µg/mL. 2 columns were coated with either BSA or HER2 and 4 columns were coated with DLL1. After washing and blocking as described in section 2.8.3, the following primary Abs were added: 1) to BSA coated wells: sheep anti-human serum albumin Ab (Abcam®) – cross reactive with BSA –, 18,17 µg/mL; 2) to HER2 coated wells: mouse anti-rhesus HER2 Ab (Sino Biological), 0,5 µg/mL; 3) to DLL1 coated wells (2 columns): mouse anti-DLL1 Ab (Sino Biological), 0,5 µg/mL; 4) to DLL1 coated wells (2 columns): mouse anti-histidine Ab, 1,4 µg/mL. Primary Ab incubated at RT for 1h25'. Washing was repeated and the following detection Abs at 2,5 µg/mL were added: 1) BSA coated wells: donkey anti-sheep IgG-HRP Ab (Santa Cruz Biotechnology®); 2-4) HER2 and DLL1 coated wells: anti-mouse IgG-HRP Ab (Sigma-Aldrich®). Following procedures were performed as described in section 2.8.3.

# 3. Results

## 3.1 Construction of two immune Ab Fab fragment libraries (κ and λ)

The biological material for the generation of the two Fab fragment immune libraries (κ and λ) was obtained at CUF Descobertas Hospital from peripheral blood from 16 breast cancer patients. The samples were drawn from female adults with ages between 48 and 84 one male with age of 71.

### 3.1.1 Light Chain and Heavy Chain Fd Fragments Construction

Total RNA was isolated from the peripheral blood mononuclear cells of breast cancer patient donors. Isolated RNA quality was evaluated by determination of its quantity, purity and integrity. Isolated RNA quantity ranged from 2,3 ng/μL to 109,3 ng/μL (Table 3.1). A260/A280 ratio, which is indicative of purity, also varied from 1,3 to 1,94, but it was higher than 1,65 for most of the samples (Table 3.1). Most of the samples were therefore in an acceptable level of purity, once this ratio optimal results are between 1,7 and 1,9 (Kotlan and Glassy, 2009).

**Table 3.1 - RNA isolated from donors' peripheral blood mononuclear cells spectrophotometry results: quantification and A260/A280 ratio.**

| Sample # | Quantity (ng/μL) | A260/A280 |
|---|---|---|
| **1** | 7,6 | 1,60 |
| **2** | 19,3 | 1,54 |
| **3** | 10,8 | 1,44 |
| **4** | 2,3 | 1,56 |
| **5** | 13,9 | 1,57 |
| **6** | 6,5 | 1,57 |
| **7** | 82,3 | 1,90 |
| **8** | 21,2 | 1,30 |
| **9** | 33,4 | 1,93 |
| **10** | 92,8 | 1,93 |
| **11** | 109,3 | 1,94 |
| **12** | 77,2 | 1,90 |
| **13** | 75 | 1,90 |
| **14** | 17,3 | 1,66 |
| **15** | 13,2 | 1,68 |
| **16** | 26 | 1,77 |

The integrity of the extracted samples was also analyzed by running an agarose gel electrophoresis, where the presence of intact 18S and 28S ribosomal RNA species (approximately 1,9 kb and 5 kb in size, respectively) was observed. Examples of three of the samples prepared are shown in Figure 3.1.

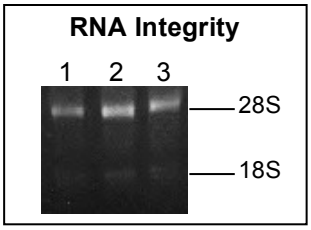


**Figure 3.1 - Integrity of RNA isolated from peripheral blood mononuclear cells from 3 donors. The two major subunits of rRNA - 18S and 28S – are detected. Lanes: 1 – Sample 14; 2 – Sample 15; 3 – Sample 16.**

After verifying that all the RNA samples passed the set quality criteria, a pool containing the same quantity of each donor total RNA was used to synthesize three different specific cDNA's. These three reactions were performed with three different primers designed to anneal with mRNA encoding either the CH1 domain of IgG, the CLκ domain or the CLλ domain. cDNA's obtained were used as first strand template to construct the HC Fd fragment, κ LC and λ LC repertoires, respectively.

Repertoires amplification was conducted in more than one step. Since the purpose was the construction of Fab libraries, both variable and constant domains (for the HC, only the first constant domain, CH1) were amplified. HC Fd and LC fragments were divided into segments (variable and constant) that were separately amplified and subsequently randomly assembled, and reamplified by overlap PCR. For the variable and CLλ segments amplification, several primer combinations were used for coverage of the different VH- and VL-gene families. Primers' design had other specifications; flanking (overlap) primers included restriction sites that were introduced in the final products extremities so that they could be cloned into the phage display vector; variable segments interior (Rv) primers included tails that were complementary to the 5' portion of the constant segment, so that both segments overlapping sequences would anneal, allowing their assembly by overlap PCR (see Table 6.1, appendix section 6.1, and Figures 2.1-2.3).

Results of the fragments construction are summarized in Table 3.2. The number of products obtained in each construction step depended on the number of primer combinations used in that stage, except for CH1 fragments, which were all constructed with the same combination of primers, but derived from different templates. Before proceeding, electrophoresis analysis of each PCR product was performed to check if bands of the correct size were obtained after each reaction. The bands of the amplified fragments with the expected sizes were gel purified to separate them from unspecific products and remaining template DNA, and hence avoid their interference in subsequent procedures.

HC cDNA synthesis with a primer specific for the IgG CH1 region allowed that the VH repertoire construction was exclusively based on V regions originated from this Ig class. VH repertoire construction comprised three essential steps. The first step consisted in the amplification of complete VHCH1 fragments from cDNA, and the 6 different products obtained presented the expected size of 750 bp (Figure 3.2 A). However, the agarose gel also showed that amplification of two of the six fragments – VH2CH1 and particularly VH6CH1 – was less efficient. Regarding this, several attempts were made to optimize VH6CH1 amplification by testing different annealing temperatures, but none had significantly better results (data not shown). Nevertheless, the amount of DNA obtained from these products was still sufficient and thus used in the following steps. In the second step the VHJH and CH1 segments were individually amplified from the previous step products, resulting in, respectively, 24 and 6 different PCR products. Those had the expected sizes of 430 bp and 364 bp, respectively, as verified by gel electrophoresis (Figure 3.2 B). In the third step, both segments obtained from the second step were assembled by overlapping PCR, originating the final HC Fd fragments. DNA gel electrophoresis analysis showed that all overlapping products were amplified successfully and presented the correct size of approximately 730 bp. However, as shown in Figure 3.2 C, besides detection of the DNA templates some unspecific bands were also present. Increasing the annealing temperature and reducing the extension time to the minimum recommended did not

improve the results, so the bands of the expected size were gel purified. An extra step was also performed in which the 6 purified overlap fragments were used as templates for their own reamplification to allow having more DNA available for the following cloning steps.

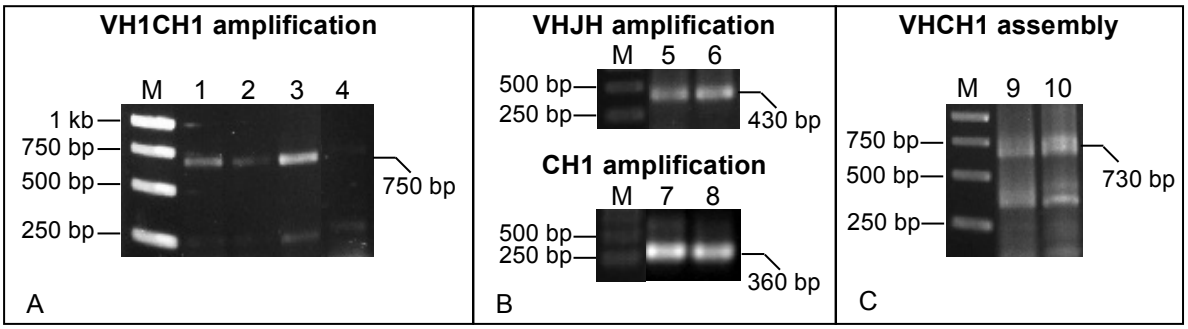


**Figure 3.2 - HC Fd fragment construction PCRs. M – Gene ruler 1 kb. A – VHCH1 amplification (1st step); Lanes: 1 – VH1CH1; 2 – VH2CH1; 3 – VH3CH1; 4 – VH6CH1, amplified with 59º C of annealing temperature. B – VHJH and CH amplification (2nd step); Lanes: 5 – VH2JH4/5; 6 – VH2JH6; 7 – 5CH1; 8 – 6CH1. C – HC Fd (VHCH1) overlap amplification (3rd step); Lanes: 9 – VH1CH1; 10 – VH2CH1.**

Both κ LC and λ LC construction was performed in two steps, first the VL and CL segments were individually amplified and then subsequently assembled. The first step resulted in 9 VLκ, 1 CLκ, 12 VLλ and 2 CLλ products. DNA gel electrophoresis results showed that all the fragments had the expected sizes – approximately 370 bp for the VLs, 337 for CLκ and 317 bp for CLλ (Figure 3.3 A and Figure 3.4 A). Moreover, the results showed that the amplification of some VL segments was less efficient, namely VLκ4 and particularly VLκ5 and VLλ10, so the amplification of those two was optimized. Better PCR results were achieved with other annealing temperatures, namely 47,1º C and 48,1º C for VLκ5 amplification (Figure 3.3 B), and 55º C, 53º C and 50,9º C for VLλ10's (Figure 3.4 B). Still, these results were not as good as for the respective remaining variants (data not shown). Considering the optimization results in the PCR assembly reaction, VLκ4CLκ and VLκ5CLκ fragments were thus constructed using an annealing temperature of 48º C. Results were equivalent to those of the remaining VLκCLκ fragments construction, 9 of which were obtained in total.

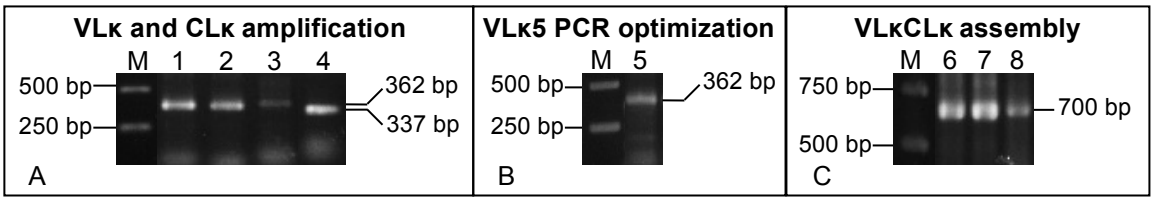


**Figure 3.3 - VLκCLκ construction PCRs. M – Gene ruler 1 kb. A – VLκ and CLκ amplification (1st step); Lanes: 1 – VLκ3c; 2 – VLκ3d; 3 – VLκ4; 4 – CLκ. B – Lanes 5: VLκ5 amplified with 48,1º C of annealing temperature (optimization). C – VLκCLκ overlap amplification (3rd step); Lanes: 6 – VLκ3cCLκ; 7 – VLκ3dCLκ; 8 – VLκ4CLκ.**

24 VLλCLλ fragments were constructed also with variable yields. DNA gel electrophoresis of the VLCL assembly products resulted in bands of 700 bp as expected (Figure 3.3 C and Figure 3.4 C).

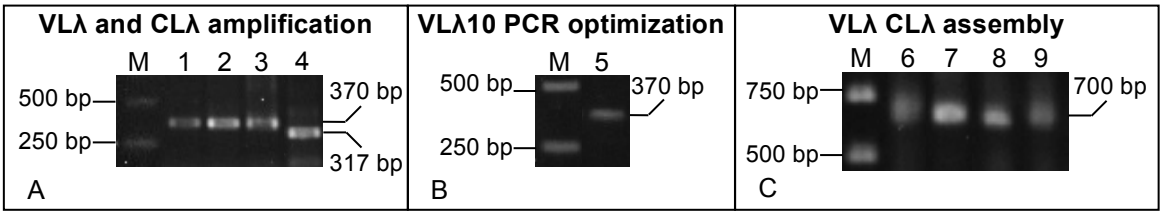


**Figure 3.4 – VLλCLλ construction PCRs. M – Gene ruler 1 kb. A – VLλ and CLλ amplification (1st step); Lanes: 1 – VLλ7/8a; 2 – VLλ7/8b; 3 – VLλ9; 4 – CLλa. B – Lanes 5: VLλ10 amplified with 50,9º C of annealing temperature (optimization). C – VLλCLλ overlap amplification (3rd step); Lanes: 6 – VLλ1aCLλa; 7 – VLλ2CLλa; 8 – VLλ7/8bCLλb; 9 – VLλ9CLλb.**

A summary of all the fragments construction results is presented below (Table 3.2).

**Table 3.2 - LC and HC Fd fragments construction: Amplification steps' results summary.**

| Repertoire | Amplification Step | Product | # of different products | Approximate size (bp) |
|---|---|---|---|---|
| **HC Fd** | 1 | VHCH1 | 6 | 750 |
| | 2 | VHJH | 24 | 430 |
| | | CH1 | 6 | 364 |
| | 3 | HC Fd (VHCH1) | 6 | 730 |
| **κ LC** | 1 | VLκ | 9 | 362 |
| | | CLκ | 1 | 337 |
| | 2 | VLκCLκ | 9 | 700 |
| **λ LC** | 1 | VLλ | 12 | 370 |
| | | CLλ | 2 | 317 |
| | 2 | VLλCLλ | 24 | 700 |

### 3.1.2 Libraries Construction

Fab libraries construction was performed by cloning the PCR generated fragments into pCOMB3XSS, by digestion with compatible restriction enzymes. Two cloning steps were performed. In the first, VLκCLκ and VLλCLλ fragments were separately cloned into the vector using the enzymes *Sac*I and *Xba*I. Subsequently, HC Fd fragments were cloned into the vectors already containing the κ or λ LC repertoires, resultant from the first cloning step, with two different enzymes – *Xho*I and *Spe*I. After these cloning steps, vectors for two libraries, VLκCLκ-VHCH1 ("κ library") and VLλCLλ-VHCH1 ("λ library"), were obtained.

Before proceeding with the construction of the large scale libraries, small scale ligation trials were performed to optimize ligation and transformation conditions, and their results are summarized in Tables 3.3 and 3.4. 1:3 and 1:5 vector:insert molar ratios were tested for LC-vector ligation and the 1:5 ratio resulted in a higher #transformants/μg DNA, namely $9{,}80 \times 10^6$ for VLκCLκ and $1{,}07 \times 10^7$ for VLλCLλ ligation. This was a positive result since the threshold was set at $1 \times 10^7$ transformants/μg DNA according to literature (Clark, 2002). Small scale trials also demonstrated that the vector self-ligation background levels were very low, only with $4 \times 10^3$ transformants/μg DNA. Since this is <0,1% of both LC-vector ligations' transformation efficiencies (Clark, 2002), one can assume that the vector was correctly digested and dephosphorylated (Figure 3.5 A). Consequently, LC large scale ligation was performed with a 1:5 vector:insert molar ratio and with the same vector batch that was used for small scale. The obtained VLκCLκ sub-library contains $5{,}17 \times 10^8$ transformants and VLλCLλ sub-library $1{,}25 \times 10^9$ transformants.

**Table 3.3 - LC small scale ligation trial results**

| Transformed DNA | #transformants/µg DNA | | | # self-ligation transformants / # 1:5 ratio vector-insert ligation transformants |
|---|---|---|---|---|
| | --- | 1:3 ratio | 1:5 ratio | |
| **pUC19** | $3,74x10^{10}$ | --- | --- | --- |
| **pCOMB3XSS self-ligation** | $4,00x10^{3}$ | --- | --- | --- |
| **pCOMB3XSS + VLκCLκ ligation** | --- | $2,56x10^{6}$ | $9,80x10^{6}$ | 0,04% |
| **pCOMB3XSS + VLλCLλ ligation** | --- | $2,35x10^{6}$ | $1,07x10^{7}$ | 0,04% |

HC Fd ligation to LC-vector small scale trials were also performed but this time only with the 1:5 ratio, once this condition was the one that allowed obtaining the best results before. The first trial ligation resulted in a transformation efficiency higher than $1x10^{7}$ transformants/µg DNA, as desired, but both LC-vectors' self-ligation background levels were >0,1%, in the 3% range. LC-vectors digestion products analysis by DNA gel electrophoresis showed that besides the band of interest (approximately 4 kb) there was another of similar size (approximately 3,5 kb) (Figure 3.5 B-C). As such, digested LC-vectors were ran in an agarose gel for longer to assure a clear separation of the bands, and ensure that only the band of interest was purified and used for the following cloning step; also, purified LC-vectors were dephosphorylated for longer times (3h instead of the initial 1h) to prevent self-ligation of the products. This resulted in a significant reduction of the self-ligation levels to the range of 0.2-0.3%, values still slightly >0,1%. Nevertheless, the self-ligation rates were still quite low when compared to the total number of transformants, and thus large scale ligations were followed and the libraries constructed. VLκCLκ-VHCH ligation transformation for the κ library construction resulted in $1,69x10^{9}$ transformants and VLλCLλ-VHCH ligation transformation for the λ library construction resulted in $4,58x10^{8}$ transformants.

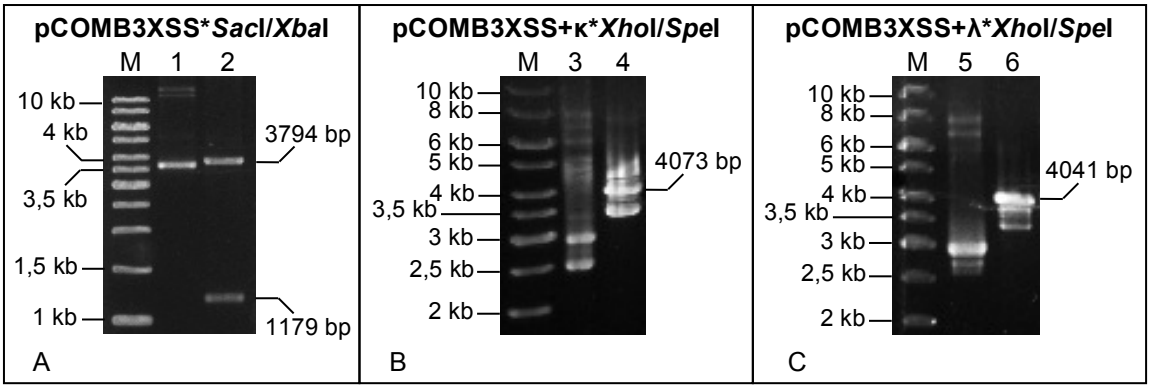


**Figure 3.5 - Vectors digestion. M – Gene ruler 1 kb. A – pCOMB3XSS. Lanes: 1 – Undigested; 2 – Digested with *Sac*I and *Xba*I enzymes. Smaller fragment corresponds to the LC stuffer. B – pCOMB3XSS+VLκCLκ. Lanes: 3 – Undigested; 4 – Digested with *Xho*I and *Spe*I enzymes. C – pCOMB3XSS+VLλCLλ. Lanes: 5 – Undigested; 6 – Digested with *Xho*I and *Spe*I enzymes.**

In both small scale and large scale ligations' transformation, a transformation with a commercial pUC19 was performed to determine the transformation efficiency of the competent TG1 cells. The value obtained was always above $1x10^{9}$ transformants/µg DNA, which indicates a high transformation efficiency of the TG1 cells that were used (Marks and Bradbury, 2004b) (values obtained in the small scale trials are indicated in Tables 3.3 and 3.4; for the large scale libraries construction it was $1,07x10^{10}$).

**Table 3.4 - HC small scale ligation trial results**

| Transformed DNA | #transformants/µg DNA | | # self-ligation transformants / # 1:5 ratio vector-insert ligation transformants | |
|---|---|---|---|---|
| | 1 | 2 | 1 | 2 |
| **pUC19** | $4,79x10^{10}$ | $1,07x10^{10}$ | --- | --- |
| **pCOMB3XSS-κ self-ligation** | $5,22x10^{5}$ | $4,56x10^{4}$ | --- | --- |
| **pCOMB3XSS-λ self-ligation** | $1,32x10^{6}$ | $1,88x10^{4}$ | --- | --- |
| **pCOMB3XSS-κ + VHCH1 ligation** | $1,92x10^{7}$ | $1,40x10^{7}$ | 2,71% | 0,33% |
| **pCOMB3XSS-λ + VHCH1 ligation** | $3,70x10^{7}$ | $1,17x10^{7}$ | 3,56% | 0,16% |

### 3.1.3 Libraries Rescue

Phage libraries were prepared by phagemid rescue from the Fab bacterial libraries with M13K07 helper phage. Rescue resulted in the production of 3,2 mL of each Fab phage library, resulting in the following concentrations: κ library – $5,25x10^{13}$ cfu/mL; λ library – $5,63x10^{13}$ cfu/mL. Each phage library was split into 16 aliquots of 200 µL, which were stored at -80ºC until further usage.

### 3.1.4 Libraries Characterization and Quality Assessment

Several tests were performed to characterize and evaluate the libraries quality. Libraries actual size was calculated considering the percentage of clones that were correctly constructed and did not contain errors. For that, the percentage of clones that contained both LC and HC inserts was firstly estimated by insert-site amplification, and then the percentage of those who were error-free was determined by analyzing their sequences. Diversity was estimated by assessing the percentage of different DNA fingerprinting patterns, V gene and CDR3 length distribution, and CDR3 sequence variability of Fab-containing clones. Results are summarized on Table 3.5. Finally, Fab-on-phage display was assessed by phage pools test by ELISA.

**Table 3.5 - Libraries characterization**

| Library | Total size[a] (# transformants) | Fab-containing transformants | Error-free Fabs | Functional library | Actual size[b] (# transformants) | Fingerprint diversity | Fingerprint + Sequencing diversity |
|---|---|---|---|---|---|---|---|
| **κ** | $1,69x10^{9}$ | 51% | 69% | 35% | $5,93x10^{8}$ | 88,6% | 93,2% |
| **λ** | $4,58x10^{8}$ | 52% | 83% | 43% | $1,98x10^{8}$ | 85,1% | 91,5% |

a – Calculated from the result of library transformants count in test plates; b – Taking into account the % of successful cloning and error-free Fab (functional library).

#### 3.1.4.1 Colony PCR and *Bst*NI Fingerprinting

The percentage of clones that contained both the LC and HC inserts was assessed by amplifying the DNA of 100 clones of each library with primers that anneal to the pCOMB3XSS vector in the regions flanking the Fab construct insertion place. The resultant PCR product should present a band of ≈1500 bp for all positive clones, corresponding to a correct construction of the Fab fragment (with both the complete HC and LC), as represented in a small sample of the clones in Figure 3.6. Considering this, and by analyzing all the clones it was estimated that 51% of κ library and 52% of λ library contained correctly constructed Fab-encoding inserts.

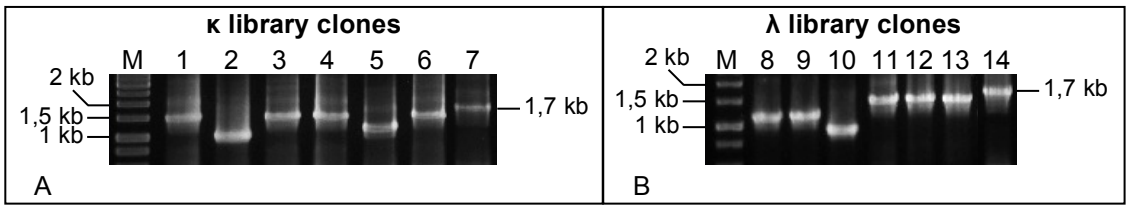


**Figure 3.6 - Example of Fab insert amplification (colony PCR). M – Gene ruler 1 kb. A – κ library clones. B – λ library clones. Lanes: 7 and 14 – pCOMB3XSS vector PCR product, 1,7 kb band (control). Clones oin lanes 1, 3, 4, 6, 11, 12 and 13 were considered positive, as their amplification resulted in a 1,5 kb band.**

44 κ and 47 λ clones that were considered positive were further analyzed by *Bst*NI fingerprinting and sequencing. *Bst*NI is a frequent-cutter enzyme and provides a DNA fingerprinting profile of the clones, which allows an estimation of the libraries' diversity. Clones digestion showed that 88,6% of the κ and 85,1% of the λ analyzed clones had different fingerprinting patterns (an example is represented in Figure 3.7).

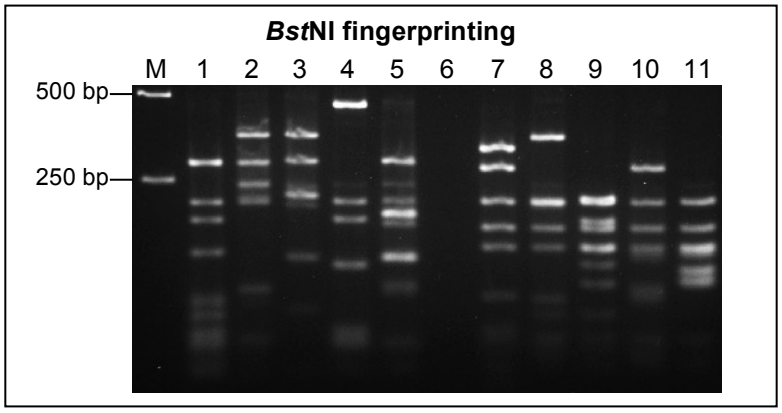


**Figure 3.7 - Example of *Bst*NI fingerprinting. M – Gene ruler 1 kb. Lanes 1-5 – κ library clones. Lanes 7-11 – λ library clones. Lane 6: empty.**

### 3.1.4.2 Sequencing

Clones that were digested with *Bst*NI were further analyzed by sequencing using the same primers as for the colony PCR. Clones' Fab insert sequencing allowed a deepest analysis of the clones diversity, identity and presence or absence of mutations, frameshifts or other cloning derived errors. Sequencing results are summarized on Table 3.6.

In total, DNA of 44 clones of the λ and 47 of the κ libraries were sent for sequencing at STAB Vida or at GATC. For 29 of the κ and 27 of the λ clones a complete sequence was obtained. For 6 of the κ and 5 of the λ clones a partial sequence was received, correspondent to only one of the chains (LC or HC Fd fragment), and 9 of the κ and 15 of the λ clones sequencing did not render any data.

**Table 3.6 - Summary of Sequencing Data**

| Library | Total sequenced | Complete sequence | 1 chain sequenced | Unmatched chains | Chains containing errors | No sequence obtained |
|---|---|---|---|---|---|---|
| **κ** | 44 | 29 | 6 | 5 | 15 | 9 |
| **λ** | 47 | 27 | 5 | 5 | 4 | 15 |

Obtained sequences were submitted to VBASE2 database analysis tool (Retter et al., 2005) that provides information about the V-genes source by aligning the sequences with known V-genes and finding the closest matches. Furthermore, it indicates when the sequences contain possible errors and finds the position and sequence of key features (like the FRs and CDRs). The results showed that most of the analyzed sequences (>91%) were comparable to the variable regions of human Igs and allowed the identification of the germ-line V genes, showing the closest nucleotide match to the clones' VH and VLκ or VLλ domains. Most of the V-gene families were represented, and VH genes were evenly distributed by all the present families in both repertoires (Figure 3.8 B and D). On the other hand, a clear dominance was seen in VL genes distribution, with a higher number of genes belonging to the VLκ3 family and VLλ1 and VLλ2 families in the respective libraries (Figure 3.8 A and C).

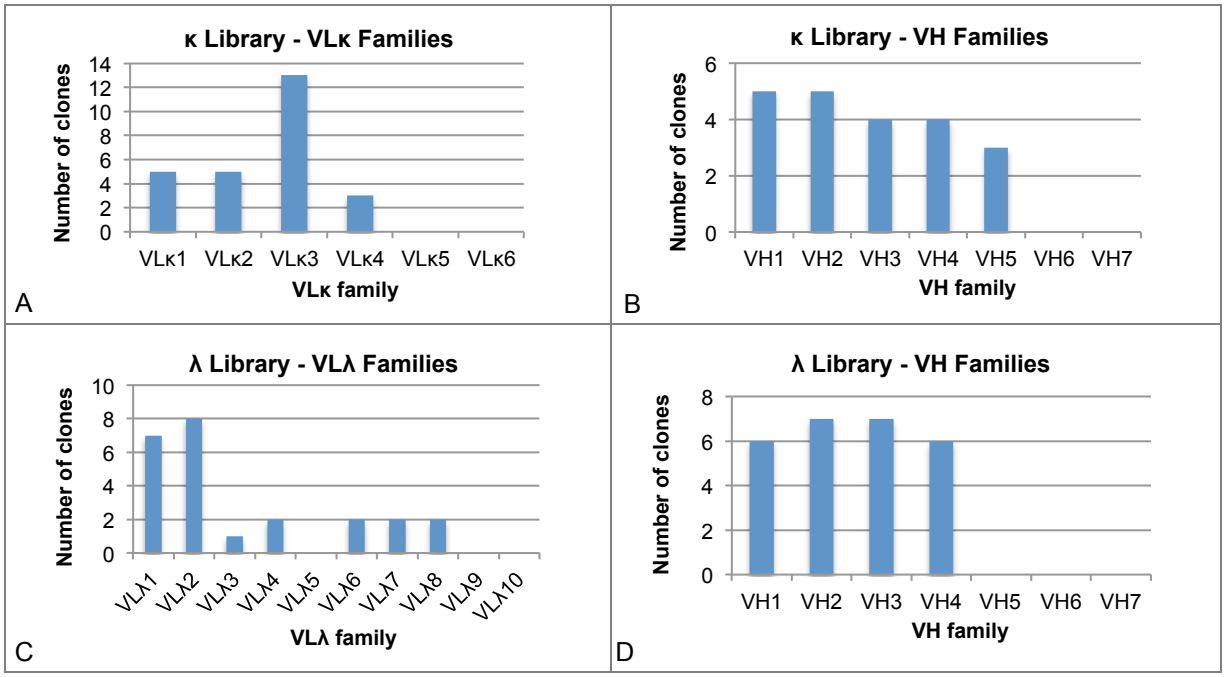


**Figure 3.8** - **V-gene families' distribution of the libraries' clones. Sequences containing errors were not considered. A – VLκ families in the κ library. B – VH families in the κ library. C – VLλ families in the λ library. D – VH families in the λ library.**

In general, there was a fair distribution of different V genes. 11 VLκ and 13 VH genes were identified in the κ library clones (Figure 3.9 A and B). IGKV3-20*01 was by far the most frequent VLκ gene, although there were 3 more genes derived from the VLκ3 family. 14 VLλ and 18 VH genes were identified in the λ library clones (Figure 3.9 C and D), being IGLV2-23*03,IGLV2-23*01 and IGLV1-44*01 the two most frequently found VLλ genes. Despite the even VH family distribution, IGHV2-5*10,IGHV2-5*02 had the higher number of VH matches in both libraries.

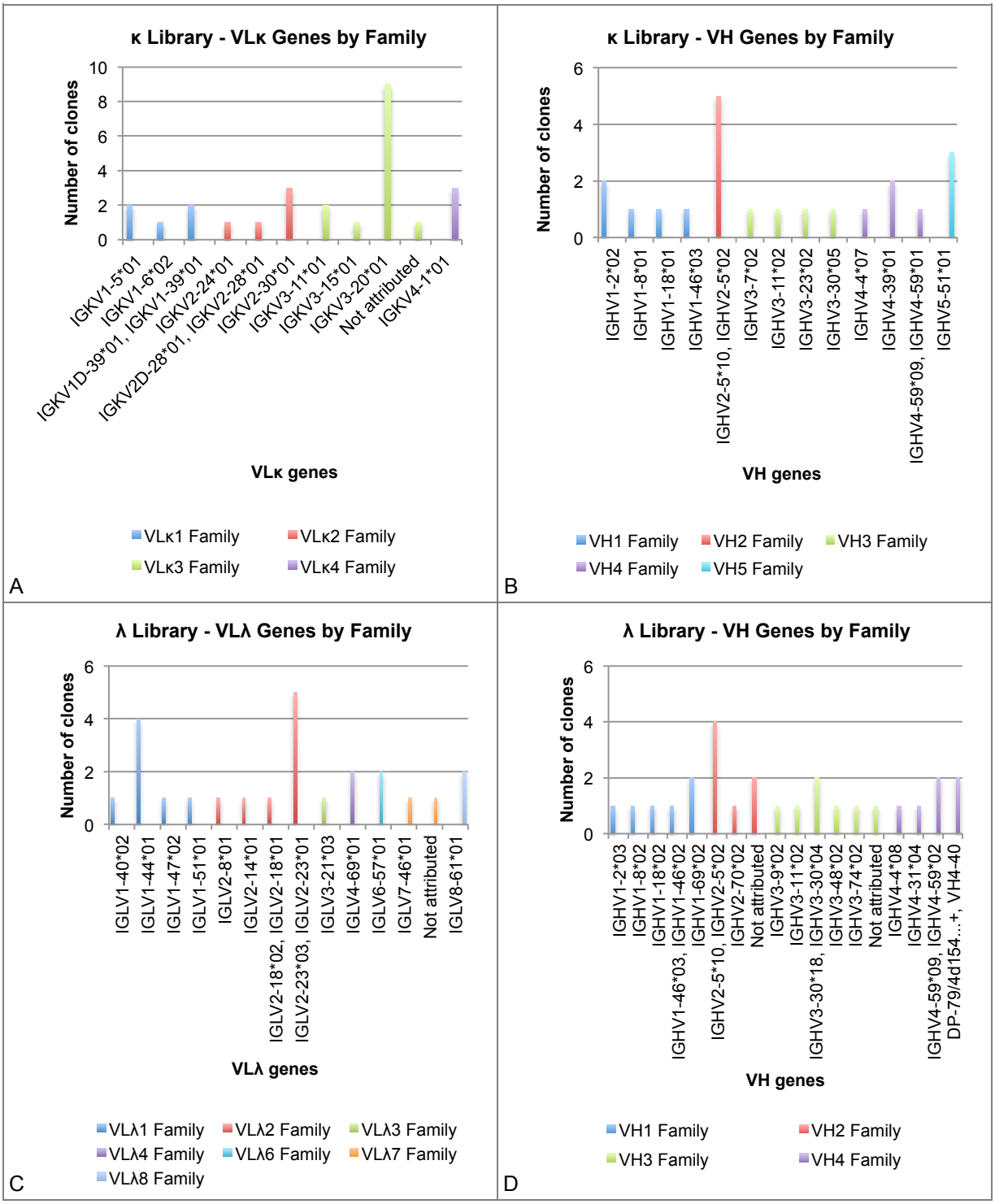


**Figure 3.9 - V gene distribution by family of the libraries' clones. Sequences containing errors were not considered. A – VLκ genes in the κ library. B – VH genes in the κ library. C – VLλ genes in the λ library. D – VH genes in the λ library.**

VBASE2 analysis tool also revealed the V region organization of the clones, including CDR3 sequence and length. VH CDR3 length had a distribution between 10 and 22 amino acid residues for the κ library, and 9 and 21 for the λ library (Figure 3.10 B and D). VL CDR3 lengths were less variable, presenting between 9 and 11 residues for the κ library, and 9 and 13 residues for the λ library, as shown in Figure 3.10 A and C. Independently of the length, CDR3 sequences were also diverse.

Excluding error-containing sequences, for the κ library 24 out of the 25 clones revealed different VL CDR3 sequences, and all 19 VH CDR3 were different; for the λ library, all 24 VL CDR3 detected were different, and 23 out of the 25 clones had distinctive VH CDR3 sequences (Tables 6.2 and 6.3, appendix section 6.2).

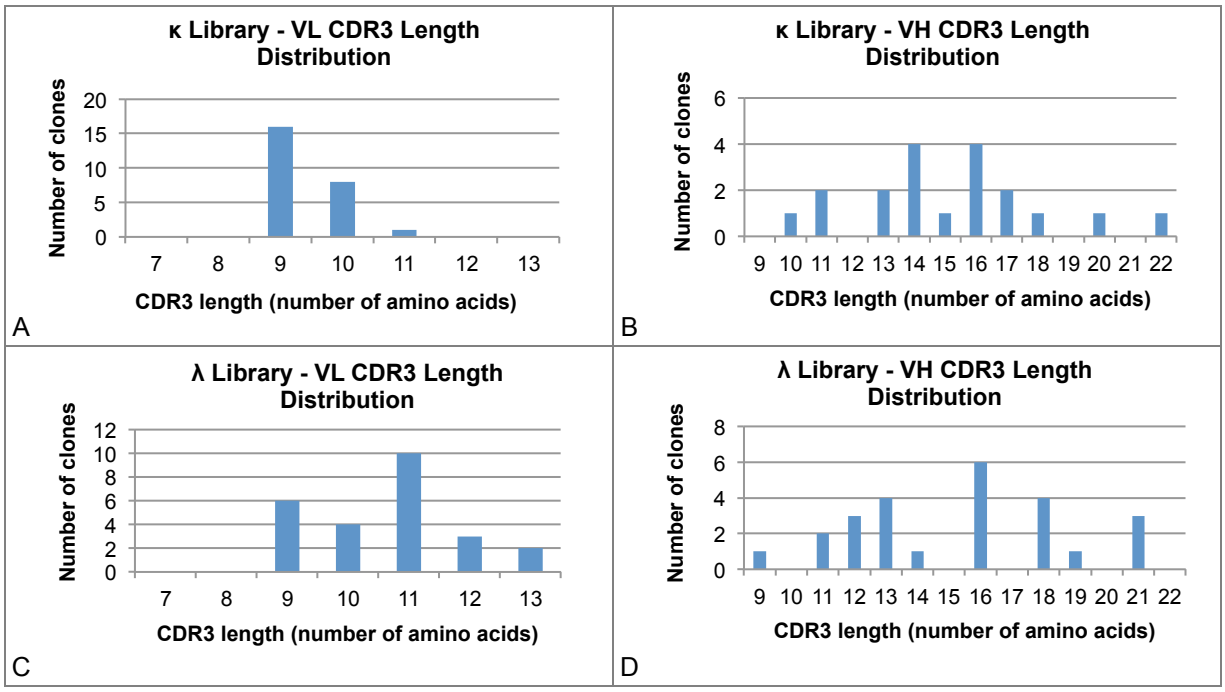


**Figure 3.10 – Libraries' CDR3 length distribution. Sequences containing errors were not considered. A – VL CDR3 in κ library. B – VH CDR3 in κ library. C – VL CDR3 in λ library. D – VH CDR3 in λ library.**

There were 5 κ and 7 λ library clone pairs that obtained similar DNA fingerprinting patterns. A comparison of the sequence analysis results of the two clones in each pair was performed when enough sequence information of both clones was available, namely for 2 κ and 3 λ library pairs. Results are shown in Table 3.7. All clones presented at least one variation, including that all the VH CDR3 sequences analyzed were different. Therefore, the clones were considered different despite the similar fingerprinting pattern, and only the clones that had the same pattern and no sequence information were regarded as possibly identical for the final library diversity assessment. Hence with this data the diversity of the libraries was higher than the previously estimated based only on DNA fingerprint (Table 3.5).

**Table 3.7 - Sequencing results of clones that obtained identical *Bst*NI fingerprinting patterns**

| Library | Clone Pair | VL gene family | VL gene | VH gene family | VH gene | VL CDR3 sequence | VH CDR3 sequence |
|---|---|---|---|---|---|---|---|
| κ | K1/K8 | Common IGKV3 | Common IGKV3-20*01 | Different IGHV1/IGHV5 | Different | Same | Different |
| | K10/K22 | Common IGKV3 | Different IGKV3-11*01/ IGKV3-20*01 | Different IGHV2/IGHV5 | Different | Different | Different |
| λ | L4/L8 | Common IGLV7 | Different Not attributed/ IGLV7-46*01 | Different IGHV3/IGHV2 | Different | Different | Different |
| | L5/L15 | Different IGLV1/IGLV2 | Different | Different IGHV2/IGHV3 | Different | Different | Different |
| | L10/L28 | Common IGLV2 | Common IGLV2-23*01 | Common IGHV1 | Different IGHV1-2*02/ IGHV1-18*01 | Unknown | Different |

### 3.1.4.3 Fab Display

The libraries Fab display was assessed by titration of the phage pools by Fab-on-Phage ELISA. Two assays were performed where phage quantity and Fab display were evaluated, the first with an anti-M13 Ab and the last with an anti-Fab specific Ab. In a second assay the specificity for the his-tag, present in the C-terminal of the Fab, was also tested using an Ab against Histidine (for His-tag detection), and to verify the correct conformation of the LC, Abs against κ or λ LC were also used (respectively). Results are shown in Figure 3.11.

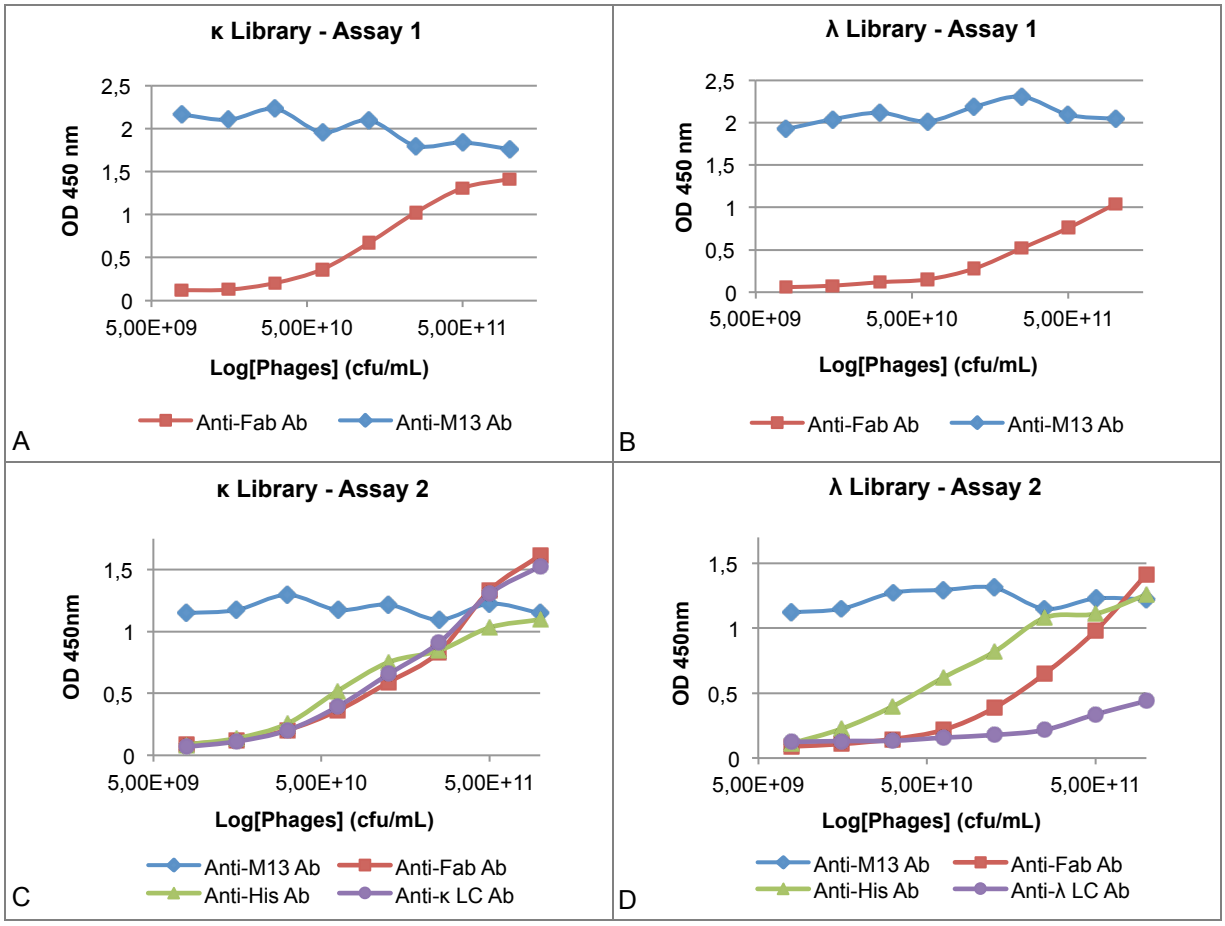


**Figure 3.11 - Fab-on-Phage ELISA of unselected libraries' phage pools. Values presented are a mean of the respective duplicates and have the background signal subtracted. A – κ library, assay 1. B – λ library, assay 1. C – κ library, assay 2. D – λ library, assay 2.**

In the first assay, the signal obtained for phage detection was very strong and saturated quickly, so no dose response was possible to be observed. Therefore, in a second assay the plates were coated with half the concentration (2,5 µg/mL) of anti-M13 Ab and the ELISA protocol followed was the same. The results were however identical, indicating that phages were still captured in very high quantities. For a more thorough assessment of the number of phage particles the ELISA protocol must be further optimized. For display assessment, signal in anti-Fab Ab coated wells had identical results in both individual experiments. All Fab display tests resulted in dose responsive signals that were stronger in the highest concentrations tested, demonstrating that both libraries were correctly constructed, since phages were displaying the Fabs correctly. Anti-Fab Ab signal for κ library was on average 1,5 fold higher than for the λ library. The difference was more pronounced for the signal

obtained with the anti-LC Abs, which was 3 to 4 fold higher for the κ library highest dilutions, but this might have been caused by the use of a different coating Ab for each assay. On the other hand, anti-his Ab signal was slightly higher for λ library, but the difference was mostly below 1,3 fold, including in the highest dilutions. Samples were analysed 2 fold each, and the signals detected for the duplicates were identical, and no significant background signal was found in uncoated wells.

## 3.2 Libraries Panning

For validation, the two constructed Fab phage display immune libraries (κ and λ) were initially selected against a commercial model antigen – BSA, which usually is highly immunogenic (Hayworth, 2014). Furthermore, and since the goal is to generate new potential cancer therapeutic Abs, selection was also pursued using as an antigen a validated breast cancer biomarker – HER2. Finally, and since the aim of this thesis is to generate function blocking Abs for Notch1 ligands, in order to inhibit Notch1 signaling pathway, selection against DLL1 was performed once this target is a known Notch1 receptor ligand.

Solid phase panning using 96-well Nunc Maxisorp plates was followed to select against all the three antigens assayed. Three rounds of selection were carried out in each case, with amplification of phages after all rounds of panning. The size and diversity of the libraries may imply that there are relatively few copies of each individual displayed Ab, as such the theoretical number of phages used in the 1$^{st}$ round was 1x10$^{12}$. Considering the size of each library, the number of phages used per selection ensures that each phage of the κ library is represented at least ≈590-fold, and each phage of λ library ≈2180-fold. As a depletion step uncoated wells were used to eliminate the unspecific plastic phage binders. Also, to increase the probability of interaction between specific phage binders and the target, in the 1$^{st}$ round of selection the phages were incubated twice consecutively with the antigen.

For each round, the total amount of phages added and recovered was measured and this allowed calculating the titer of the input and output phages, and determining the percentage of input recovery and enrichment values.

The amplified phage pools resultant from the output were analyzed by Fab-on-phage ELISA to check if target-specific phages were being recovered and enriched. Amplified phages concentration was not determined, so the ELISA results are used as a qualitative indication. After three rounds of selection, individual clones were picked and analyzed by Fab-on-phage ELISA. Moreover, clones that showed a good Fab display, reactivity for the respective antigen and no background reaction were considered positive, and thus further tested by sFab ELISA. The sFab ELISA data allows checking the clone ability to produce reactive soluble Fab fragments. In cases where samples were analysed 2-fold each, the signals detected for the duplicates were identical.

Positive clones were further characterized. Phagemid DNA of the hit clones was extracted and the presence of the Fab insert was evaluated by PCR using the primers that anneal to the pCOMB3XSS vector in the regions flanking the Fab construct insertion place (the same that were used for the libraries size characterization). Phagemid DNA was also digested with *Sac*I and *Spe*I and the restriction pattern analyzed.

### 3.2.1 BSA

Selection was carried out using 20 µg/mL of the BSA antigen molecule. Input and output phage titers, the percentage of input recovery and enrichment values are indicated in Table 3.8.

**Table 3.8 - BSA panning**

| Target | Panning round | | Phage input (cfu) | Phage output (cfu) | % of input phage recovery ($x10^{-4}$) | Enrichment* |
|---|---|---|---|---|---|---|
| BSA | κ | 1 | $1{,}73x10^{12}$ | $2{,}85x10^{6}$ | 1,65 | --- |
| | | 2 | $8{,}44x10^{12}$ | $5{,}00x10^{2}$** | 0,0001** | $3{,}60x10^{-5}$x** |
| | | 3 | $4{,}80x10^{11}$ | $5{,}00x10^{5}$ | 1,04 | 17583,33x** |
| | λ | 1 | $1{,}27x10^{13}$ | $3{,}76x10^{4}$ | 0,003 | --- |
| | | 2 | $2{,}46x10^{11}$ | $2{,}65x10^{4}$ | 0,11 | 36,27x |
| | | 3 | $2{,}33x10^{11}$ | $4{,}98x10^{6}$ | 21,35 | 198,21x |

*-fold increase in recovered phage, compared to the previous round of panning

** The data obtained is not within the acceptance quality criteria (Rader, 2012) (see discussion section)

In the 1st round of selection, the κ library input calculated value ($1{,}73x10^{12}$) was close to the expected, once $1x10^{12}$ phages had been initially added for selection. For the λ library, however, it was considerably higher (in the $10^{13}$ range – $1{,}27x10^{13}$), which indicates that more phages were added than intended, possibly due to a dilution error. The number of phages added to selection was only predetermined in the 1st round. For the subsequent rounds a defined volume of the previous round amplified phages was added instead. In the 2nd and 3rd rounds, the results obtained for the titration of the input were in the range of $10^{11}$-$10^{13}$ phages for κ library, which means they were in the range of the 1st round's. For λ library, 2nd and 3rd round values were very similar (~$2x10^{11}$) but differed slightly from that of the respective 1st round, nevertheless the results are in the typical expected range of $10^{11}$-$10^{12}$ total phages (Rader, 2012).

For the κ library the output titration for the 1st round presented the value of $2{,}85x10^{6}$ phages, which decreased dramatically in the 2nd round, to $5{,}00x10^{2}$. This output value was lower than expected, once the expected value usually ranges from $10^{4}$ to $10^{7}$ (Rader, 2012). Nevertheless, a decrease in the output is typically seen from the 1st to the 2nd round and it could have been exacerbated by a possible tittering error. In the 3rd round the output increased to $5{,}00x10^{5}$. Due to the significant difference in the 2nd and 3rd round output values, the calculated 3rd round enrichment value was extremely high. Assuming that, as referred, the 2nd round output was miscalculated, this enrichment value should not be taken into consideration for further discussion.

Output titration of the λ library selection showed a slight decrease from the 1st to the 2nd round, from $3{,}76x10^{4}$ to $2{,}65x10^{4}$, which was much less significant than for the κ library. In the 3rd round, the output increased to $4{,}98x10^{6}$, which can be translated in a significant enrichment value.

#### 3.2.1.1 BSA Selected Phage Pools Reactivity

Since this was the first target to be assayed, for BSA, both eluted and amplified phage pools were analyzed by Fab-on-phage ELISA. Eluted phages concentration was calculated based on the phage output values. In addition to an anti-Fab Ab, Abs against κ or λ LC (depending on the library) were also used to evaluate Fab expression.

When analyzing the Fab-on-phage ELISA results for the eluted phages of the 1st and 2nd rounds of selection no signal was detected neither for Fab display nor for BSA reactivity, for both libraries. In an attempt to obtain better signals, revelation reactions with TMB were allowed to take longer but with no significant improvement (data not shown). This was probably due to the fact that the eluted phages concentration was not high enough to reach the assay detection limit.

As such, amplified phages were analyzed for all rounds of selection. In the case of κ library (Figure 3.12), the results obtained showed that there was a dose-dependent behavior for Fab presence as can be seen in the curves obtained for both the anti-Fab and anti-LC assay. Moreover, the values obtained for both Abs were in the same range. However, no significant signal or response was detected for BSA reactivity, indicating there was not a BSA-specific phage enrichment in the selection strategy followed.

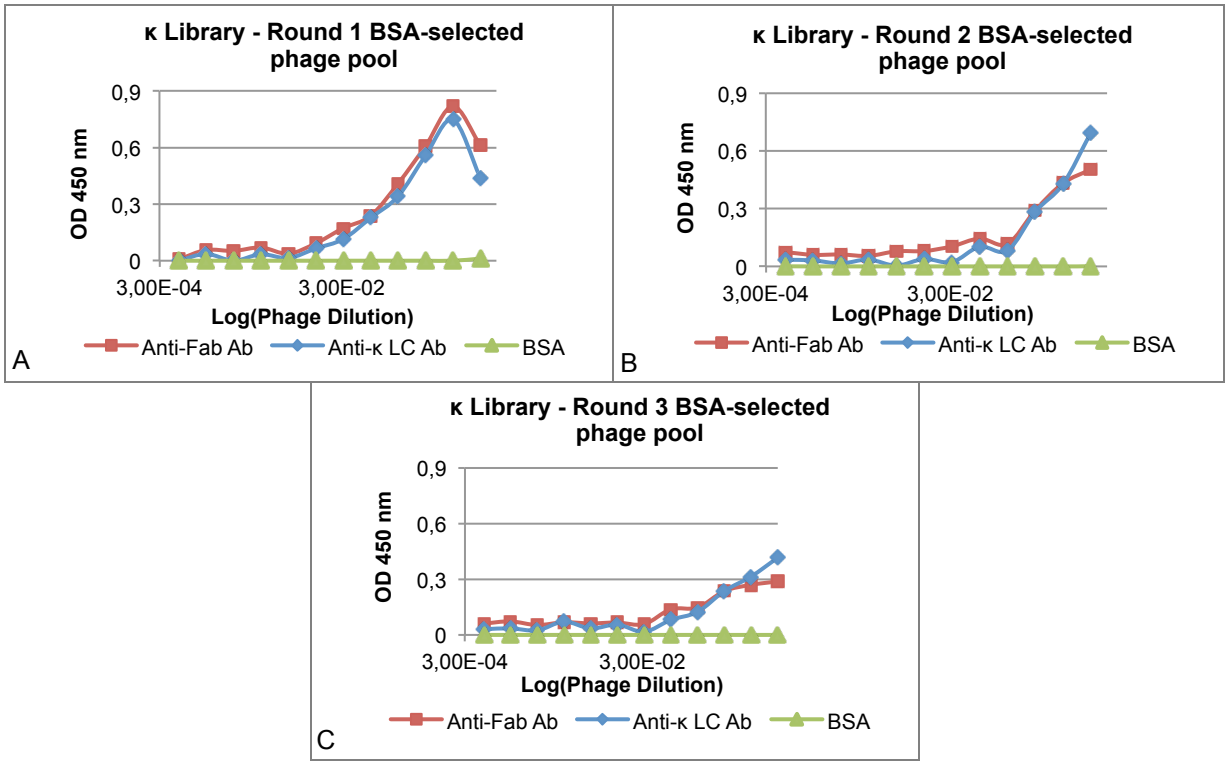


**Figure 3.12** - **Fab-on-Phage ELISA of amplified phage pools resultant from κ library selection against BSA. Values presented are a mean of the respective duplicates and have the background signal subtracted. A – 1st round of selection. B – 2nd round of selection. C – 3rd round of selection**

When analyzing the results for the Fab-on-phage ELISA for the λ library amplified phages there was a dose response towards the anti-Fab and the anti-LC Abs, indicating correct display of the Fab's and respective LC. In regards to BSA reactivity the results obtained show that there was a specific dose-responsive signal, which increased from round to round suggesting the enrichment of BSA-specific phages. Moreover the ELISA signals obtained for the highest concentrations tested were at least 3,5-fold over the background (Figure 3.13).

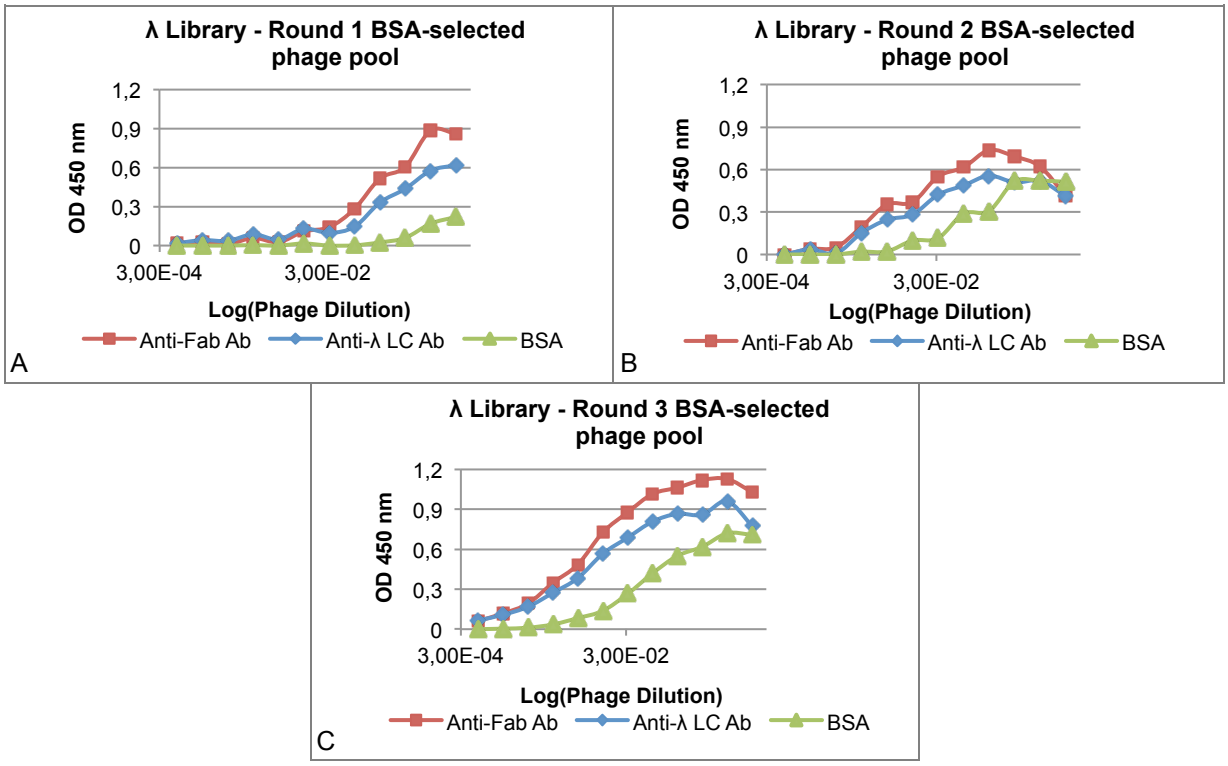


**Figure 3.13 - Fab-on-Phage ELISA of amplified phage pools resultant from λ library selection against BSA. Values presented are a mean of the respective duplicates and have the background signal subtracted. A – 1st round of selection. B – 2nd round of selection. C – 3rd round of selection.**

Considering the previous results, only the λ library was used to continue with our studies, and thus this was used to select against the other two antigens.

### 3.2.1.2 BSA Selected Individual Clones Reactivity

Since the κ library selected phage pool analysis showed that it was not reactive for BSA, only selected λ clones were further tested. 10 individual clones were randomly picked from the 3rd round output of the λ library selection and analyzed by Fab-on-phage ELISA. Similarly to before, Fab expression was also evaluated using Abs against Fab and λ LC.

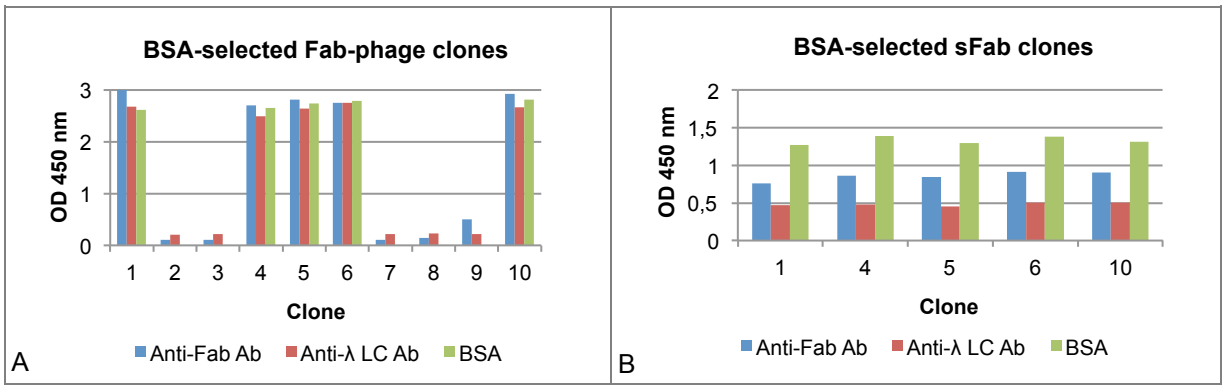


**Figure 3.14 - ELISA of λ library individual clones selected against BSA after 3 rounds of panning. Values presented are a mean of the respective duplicates and have the background signal subtracted. A – Fab-on-Phage ELISA. B – sFab ELISA.**

5 out of 10 (50%) of the BSA tested clones (B1, B4, B5, B6 and B10) were considered positive in both Fab-on-phage and sFab ELISAs, since the ELISA signals obtained for BSA reactivity were about 40-fold higher than the background for the Fab-on-phage assay, and over 20-fold higher than the background for the sFab assay (Figure 3.14).

#### 3.2.1.3 BSA Positive Clones Characterization

The clones that gave a positive signal upon Fab-on-Phage and sFab ELISA analysis were further characterized. This was done by an initial PCR amplification of the Fab insert to confirm the size and by a *Sac*I/*Spe*I digestion of the extracted phagemid DNA. It was expected that the PCR products presented a band of ≈1500 bp, however by observing the DNA agarose gel (Figure 3.15) it can be seen that the bands of the amplified products have a higher molecular weight of ≈1800-1900 bp. Moreover, B10 clone amplification was less efficient than the remaining, resulting in a faint band as seen on the gel. In regards to the phagemid digest, the pattern obtained showed that the digestion products were smaller than expected (about 1400 bp instead of 1500 bps), so apparently the clones selected and analyzed might not have the correct Fab sequence.

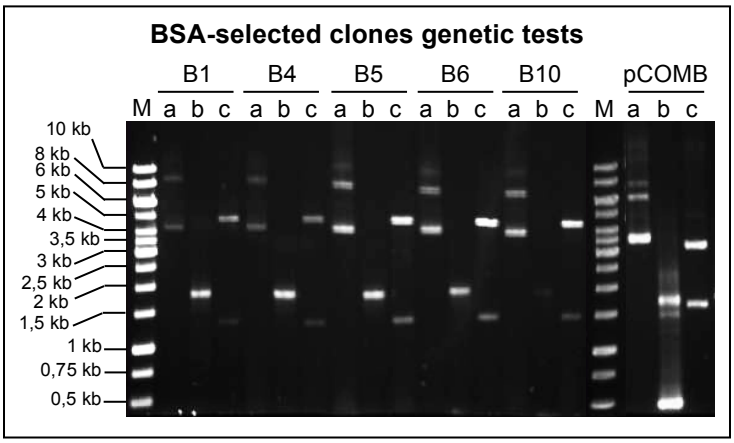


**Figure 3.15 - BSA-reactive clones genetic characterization. Lanes: M – Gene ruler 1 kb; a – Undigested phagemid DNA; b – PCR amplification of the Fab insert from phagemid DNA; c – Phagemid DNA digestion with *Sac*I/*Spe*I.**

In order to better understand the previously described results, phagemid DNA from the 5 clones was sent to both Stab Vida and GATC for sequencing. However, after several attempts, the results were not satisfactory since the data delivered always presented a very short sequence or none at all. To try and overcome this issue sequencing was also tried with the Fab insert amplification products, but this also failed. Thus no clear conclusion could be drawn in regards to the BSA clones that were identified by both Fab-on-phage and sFab ELISAs. In the future, a more thorough analysis must be carried out to allow clarifying the obtained results.

### 3.2.2 HER2

λ library selection was carried out using 15 µg/mL of the HER2 (Fc tagged) antigen molecule. An extra depletion step was performed with a human IgG1-coated well, after the plastic binders depletion, to eliminate any Fc binders. Input and output phage titers, the percentage of input recovery and enrichment values are indicated in Table 3.9.

**Table 3.9 - HER2 panning**

| Target | Panning Round | Phage input (cfu) | Phage output (cfu) | % of input phage recovery ($x10^{-4}$) | Enrichment* |
|---|---|---|---|---|---|
| HER2 | 1 | $3,77x10^{11}$ | $4,82x10^{5}$ | 1,28 | --- |
| | 2 | $2,66x10^{11}$ | $6,63x10^{3}$ | 0,02 | 0,02x |
| | 3 | $1,37x10^{11}$ | $1,24x10^{6}$ | 9,06 | 364,46x |

*-fold increase in recovered phage, compared to the previous round of panning

Input phage value was stable between rounds, in the $10^{11}$ range. Considering that two depletions were performed before selection, the number of phages that were effectively selected is in accordance to the number of phages theoretically added in the 1st round ($1x10^{12}$).

The 1st round, output titration value was $4,82x10^{5}$ phages, decreasing to $6,63x10^{3}$ in the 2nd round, and increasing to $1,24x10^{6}$ in the 3rd, which can be translated in a significant enrichment.

### 3.2.2.1 HER2 Selected Phage Pools Reactivity

For the HER2 target, polyclonal ELISA analysis was performed for the 2nd and 3rd rounds using the amplified phage pools. Due to the fact that the HER2 molecule used for this panning presented an Fc fusion, the specificity for this moiety of the target was also assessed. Results are shown in Figure 3.16.

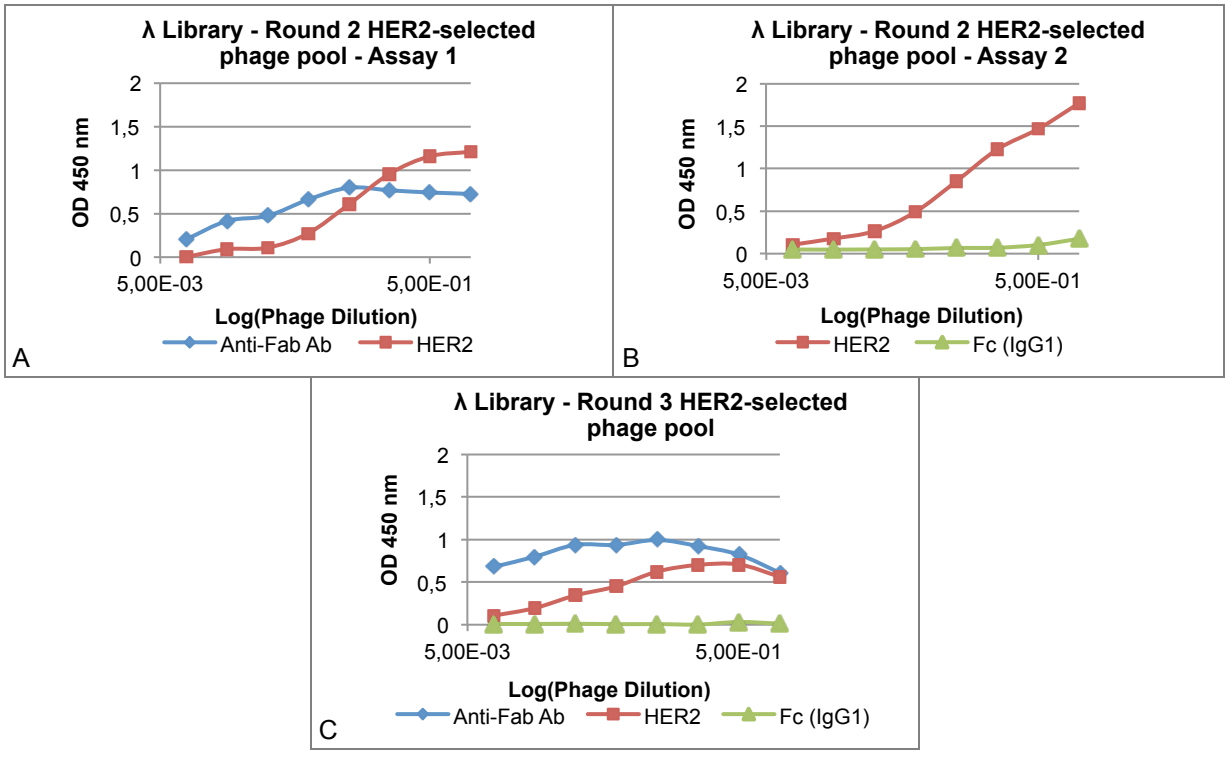


**Figure 3.16 - Fab-on-Phage ELISA of amplified phage pools resultant from λ library selection against HER2. Values presented are a mean of the respective duplicates and have the background signal subtracted. Exception is the values for assay 2 for the 2nd round of selection, in which the background signal was not assessed. A – 2nd round of selection, assay 1. B – 2nd round of selection, assay 2. C – 3rd round of selection.**

HER2 selected phages' assays demonstrated that there was a dose response signal for HER2 reactivity in the 2nd and 3rd rounds and the values obtained for the highest concentrations tested

were at least 4-fold over the background. In the 3rd round there is an abnormal result for the highest concentration of phages, but since all the dilutions follow the same dose response tendency we will disregard this value for further discussion. Fab display signal was dose responsive in the 2nd round for the lowest dilutions, but then it stabilized. For the 3rd round, the Fab display signal was not as expected; typically, the signal increased along with the phage concentration, but when analyzing the last three values the data shows a contrary behavior. For clarification of this matter another experiment will have to be performed. Nevertheless, the signal values obtained were at least 2-fold over the background and therefore are indicative of Fab presence. In all cases, no signal was obtained for Fc reactivity, as expected.

#### 3.2.2.2 HER2 Selected Individual Clones Reactivity

20 individual clones were randomly picked from the 3rd round output of the λ library selection and analyzed by Fab-on-phage ELISA. 4 out of 20 (20%) HER2 tested clones (H3, H9, H19 and H20) were considered positive in both Fab-on-phage and sFab ELISA (Figure 3.17) by showing strong signals for Fab display and HER2 reactivity, and also no signal for Fc affinity. Signal for HER2 reactivity was over 20-fold higher than the background for the Fab-on-phage assay, and over 17-fold higher than the background for the sFab assay. Fc affinity was not screened in sFab ELISA because the detection Ab used is specific for IgG Fab and this would detect the IgG1 utilized as Fc target; nevertheless, since the result was negative in the Fab-on-phage format the same would be expected and thus extrapolated for the sFab format.

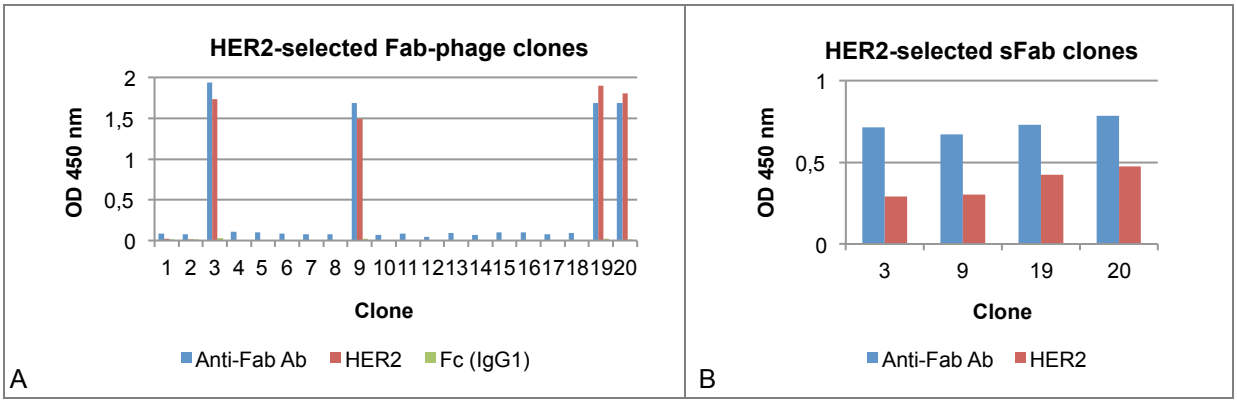


**Figure 3.17 - ELISA of λ library individual clones selected against HER2 after 3 rounds of panning. Values presented are a mean of the respective duplicates and have the background signal subtracted. A – Fab-on-Phage ELISA. B – sFab ELISA.**

#### 3.2.2.3 HER2 Positive Clones Characterization

The positive hits for HER2 identified in both the Fab-on-phage and sFab ELISA were further analyzed. As such, a PCR reaction was pursued for HER2 positive clones from extracted phagemid DNA to determine the size of the Fab insert. Clones DNA PCR had variable yields and as one can observe in Figure 3.18 H20 clone amplification was the most efficient, while H9 clone apparently did not work at all. The PCR products' bands were larger than expected, with ≈1700-1800 bp instead of the 1500 bp determined for the Fab insert. Clones' phagemid DNA was also analyzed by *Sac*I/*Spe*I digestion. By observing the gel it was visible that instead of the two bands of ≈3500 bp and 1500 bp

for pCOMB3XSS and Fab insert, the digestion products of the clones presented two bands with different molecular weights, ≈4500 bp for the vector moiety and 1300 bp-1400 bp for the insert. Thus it is clear that there was an error in the construction of these selected clones, and no correct Fab sequence is present.

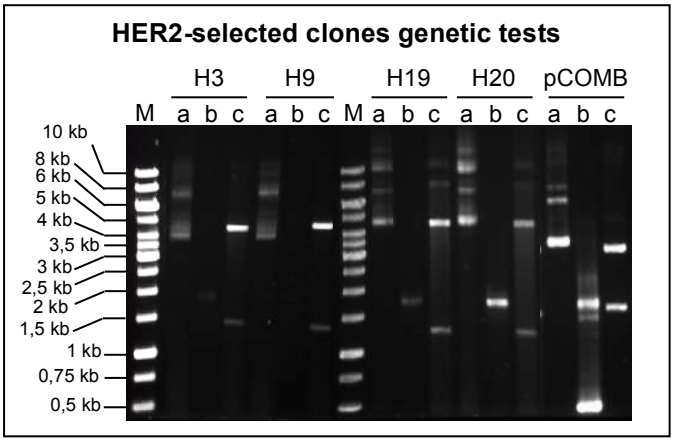


**Figure 3.18 - HER2-reactive clones genetic characterization. Lanes: M – Gene ruler 1 kb; a – Undigested phagemid DNA; b – PCR amplification of the Fab insert from phagemid DNA; c – Phagemid DNA digestion with *Sac*I/*Spe*I.**

DNA of the 4 clones was sent for sequencing to obtain more specific and clarifying data. Similarly to what happened to the BSA clones, despite the several attempts, HER2 clones sequencing was not possible with the phagemids DNA neither with the Fab insert amplification products. Very small or no sequences at all were obtained, and according to the companies' these issues could be linked to low amount and purity of the samples. Thus for a new assessment, new DNA extraction would have to be carried out.

### 3.2.3 DLL1

The λ library panning strategy was carried out using 15 µg/mL of the DLL1 antigen molecule in all rounds. Input and output phage titers, the percentage of input recovery and enrichment values are indicated in in Table 3.10.

Table 3.10 - DLL1 panning.

| Target | Panning Round | Phage input (cfu) | Phage output (cfu) | % of input phage recovery ($x10^{-4}$) | Enrichment* |
|---|---|---|---|---|---|
| **DLL1** | 1 | $5,17x10^{11}$ | $1,01x10^{6}$ | 1,96 | --- |
| | 2 | $3,10x10^{11}$ | $1,43x10^{4}$ | 0,05 | 0,02x |
| | 3 | $2,93x10^{11}$ | $2,62x10^{6}$ | 8,97 | 194,90x |

*-fold increase in recovered phage, compared to the previous round of panning

Input phage titration results obtained were in the $10^{11}$ range for the three rounds, values that are the vicinity of the number of phages used in the 1st round, and the slight decrease is probably due to the initial depletion of the plastic binders.

The output titration decreased from $1,01x10^{6}$ phages to $1,43x10^{4}$ from the 1st to the 2nd round, and the results obtained were in the usual expected range. In the 3rd round, the titration of the output

phages gave the value of 2,62x10$^{6}$, superior to the one obtained in the 2$^{nd}$ round, indicating that there was a significant enrichment from one round to another.

### 3.2.3.1 DLL1 Selected Phage Pools Reactivity

Polyclonal ELISA analysis was performed for the 2$^{nd}$ and 3$^{rd}$ rounds of the selection against DLL1 using the amplified phage pools. The results demonstrated that there was signal for both Fab display and DLL1 reactivity with 2$^{nd}$ round phages (Figure 3.19 A). Dose response was observed, especially for DLL1 affinity, for which the values obtained for the highest concentrations tested were at least 5,5-fold over the background. However, the same was not verified with 3$^{rd}$ round phages. Although Fab presence and DLL1 reactivity were apparent, no dose response pattern could be detected, thus the assay was repeated with more dilutions. The results obtained were similar (Figure 3.19 B), which suggests that the unexpected behavior might not have derived from an error in the assay itself but could come from a problem with the amplification process. For further assessment a new amplification of the 3$^{rd}$ round phages would be recommended.

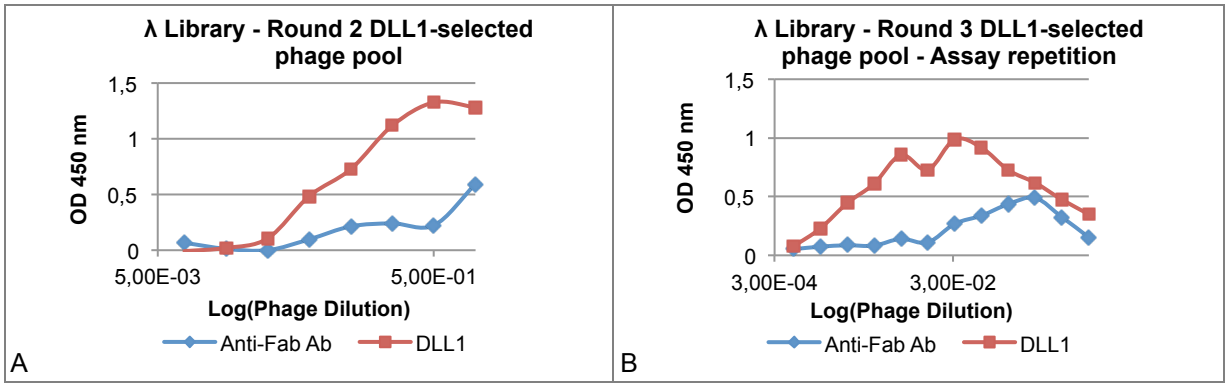


**Figure 3.19 - Fab-on-Phage ELISA of amplified phage pools resultant from λ library selection against DLL1. Values presented are a mean of the respective duplicates and have the background signal subtracted. A – 2$^{nd}$ round of selection. B – 3$^{rd}$ round of selection, assay repetition.**

### 3.2.3.2 DLL1 Selected Individual Clones Reactivity

Although the results of the 3$^{rd}$ round amplified phage pools tests were not ideal, they still suggested reactivity for DLL1, so 20 individual clones were randomly picked from the 3$^{rd}$ round output and analyzed by Fab-on-phage ELISA. The results analysis was not straightforward, since none of the tested clones presented a strong signal for Fab display and DLL1 reactivity (Figure 3.20 A). Yet, the 3 out of the 20 clones tested by Fab-on-phage ELISA that showed a higher signal (clones D2, D3 and D19) and thus stronger DLL1 reactivity (at least 2,3-fold over the background) were also tested in the sFab ELISA format. In a 1$^{st}$ sFab assay, D2 and D3 clones did not show a significant signal neither for Fab display nor for DLL1 reactivity. In the 2$^{nd}$ sFab assay, D2 and D3 clones were retested and the D19 clone tested for the first time; clones' histidine reactivity was also assessed, once the DLL1 protein used for selection was His-tagged. The clones showed inconsistent responses and very low signals were obtained towards DLL1 binding (Figure 3.20 B). Nevertheless, and due to the data obtained for these clones in the Fab-on-phage format further characterization was pursued to allow identifying possible DLL1 specific candidates.

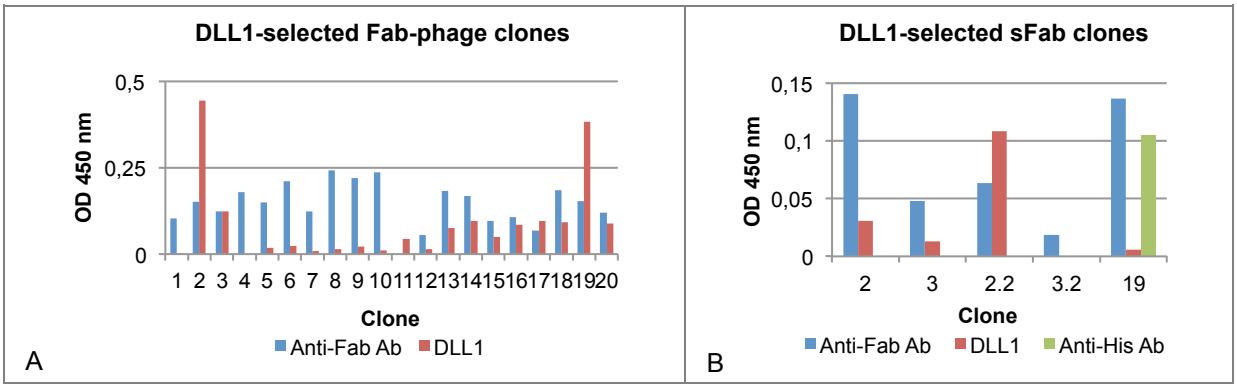


**Figure 3.20 - ELISA of λ library individual clones selected against DLL1 after 3 rounds of panning. Values presented are a mean of the respective duplicates and have the background signal subtracted. Exception is the values for assay 2 for sFab clones 2 and 3, in which no duplicates were performed. A – Fab-on-Phage ELISA. B – sFab ELISA; 2 - clone 2, assay 1. 3 - clone 3, assay 1. 2.2 - clone 2, assay 2. 3.2 - clone 3, assay 2. 9 - clone 9, assay 2.**

### 3.2.3.3 DLL1 Clones Characterization

A PCR reaction of the three clones that gave a positive hit in the Fab-on-phage ELISA was carried out. Subsequently, the Fab insert amplification of the DLL1 positive clones was analyzed on an agarose gel and the results are shown in Figure 3.21 A. For D2 and D19 clones both PCR products resulted in a band of ≈1500 bp as expected. Contrarily, the results obtained for the D3 clone PCR amplification showed that the reaction was not efficient. Simultaneously, phagemid DNA of all three clones was subjected to a *Sac*I/*Spe*I digestion for restriction pattern analysis. By observing the gel both D2 and D19 presented the expected pattern of 1500 bp for the insert, and 3794 bp for the remaining moiety of the pCOMB3XSS vector. However, the D3 clone digestion products had totally unexpected sizes, but this was in total agreement with the results obtained for the undigested phagemid DNA run in an agarose gel. As such from this analysis clones D2 and D19 seemed to be in perfect condition to continue with further characterization.

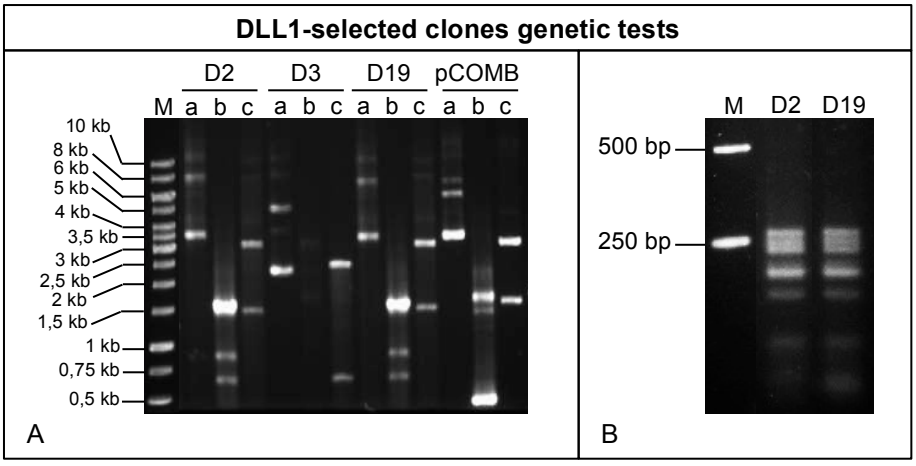


**Figure 3.21 - DLL1-reactive clones genetic characterization. M – Gene ruler 1 kb; A – Lanes: a – Undigested phagemid DNA; b – PCR amplification of the Fab insert from phagemid DNA; c – Phagemid DNA digestion with *Sac*I/*Spe*I enzymes. B - *Bst*NI digestion of D2 and D19 clones**

In order to verify if the two positive clones D2 and D9 had the same sequence an initial assessment was carried out by digestion of the PCR products with *Bst*NI. The results are shown in Figure 3.21 B, and as can be seen the clones present a similar fingerprinting pattern.

Phagemid DNA from all three clones had been initially extracted and sent for sequencing. The sequencing results obtained confirmed that D2 and D19 clones were identical. The sequence is shown in appendix section 6.3 As expected due to the results previously obtained for the D3 clone, this clone's sequence was shorter and corresponded to the LC sequence alone.

Clones' V domain characterization, obtained from VBASE2 analysis tool, is indicated in Table 3.11. D2/D19 LC gene belongs to the IGLV4 V-gene family and the HC to the IGHV3 V-gene family. LC CDR3 comprises 9 amino acid resides and HC CDR3 15, a medium size considering the lengths distribution obtained for the λ library.

**Table 3.11 - DLL1 clones' V domain characterization**

| | | Clone | |
|---|---|---|---|
| | | D2/D19 | D3 |
| **VL** | V-gene family | IGLV4 | IGLV8 |
| | V-gene | IGLV4-69*01 | IGLV8-61*01 |
| | CDR1 | SGHKNFA | CSVSTNNY |
| | CDR2 | LDSEGSH | NTN |
| | CDR3 | QTWATDGVS | VLYIHVGIWV |
| **VH** | V-gene family | IGHV3 | --- |
| | V-gene | IGHV3-33*01 | --- |
| | CDR1 | GFTFSSYG | --- |
| | CDR2 | IWYDGSNK | --- |
| | CDR3 | ARDANSWYSGNYFDH | --- |

After the genetic characterization of both D2 and D19 clones, a further analysis was pursued to verify the expression levels of the Fab's. As such, the same samples of D2 and D19 that were tested for sFab expression in the ELISA format, were analyzed by SDS-PAGE under reducing conditions. Their expression was assessed by Western Blot using an anti-Fab specific Ab. The results show that a band of ≈25 kDa was detected for both clones, correspondent to the LC and the Fd fragment in monomeric form (Figure 3.22).

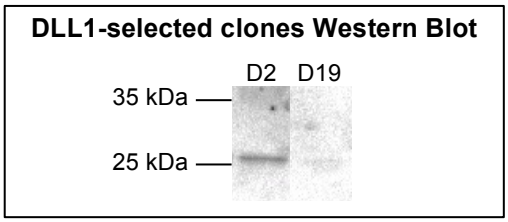


**Figure 3.22 - DLL1-reactive clones' sFab expression analysis. sFab's were run on SDS-PAGE under reducing conditions and Western Blot detection was performed with an anti-Fab Ab.**

A significant difference between the intensity of the bands of the two clones was however noticed. Since these clones are identical it would be expected to obtain bands with an identical intensity for both clones, once no variation should be detected in the expression yield. However, the

sample used for Western Blot was the same in terms of volume and not in terms of cell OD and this could support the difference reported. However, for a more clear assessment of the expression of each clone should be carried out in optimized conditions.

### 3.3 Antigens Titration

Antigens titration was performed to evaluate the panning targets quality. Duplicate samples gave identical results. As shown on Figure 3.23, BSA and HER2 titration had a dose dependent signal, suggesting these proteins are in good conditions and responding properly to the doses assayed. In the case of DLL1, the detection was performed with an anti-DLL1 Ab and an anti-His Ab because this protein was His-tagged. The ELISA signals obtained for the detection with the anti-His Ab were very low, but a small dose response effect was revealed in the higher concentrations of DLL1 tested. However, detection with the anti-DLL1 Ab had no significant signal, and this can be due to the fact that the antigen used in the panning does not comprise the whole extra-cellular domain of the protein (see discussion).

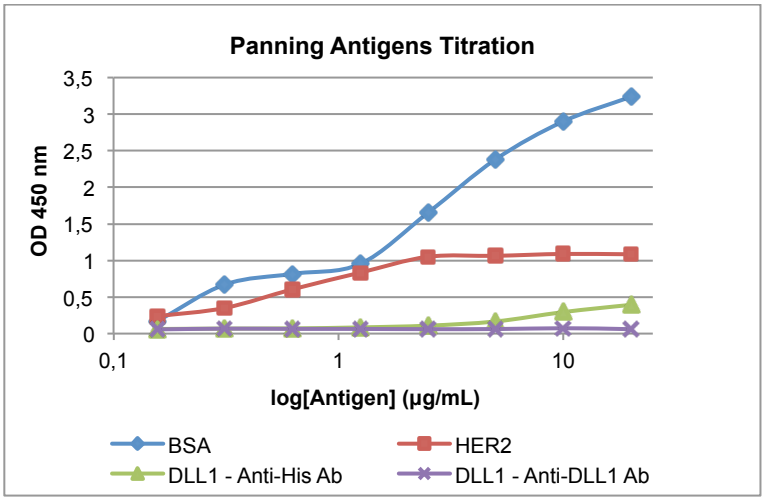


**Figure 3.23 - ELISA analysis of the antigens used for panning. Values presented are a mean of the respective duplicates and have the background signal subtracted.**

The difference in the ELISA signals obtained can also be due to the fact that all the antigens were tested simultaneously in the same plate. Although the proteins concentration was the same their detection was carried out with different specific Abs, and as such the best time for development of each reaction might not be the same for all. For a better assessment the titration of each antigen should be optimized.

# 4. Discussion

The undeniable burden and impact of breast cancer in today's world emphasize the need of persistent research in this area, including the exploration of novel therapy alternatives. Targeted therapy has an increasing importance given the present recognition of breast cancer molecular heterogeneity as a restraint to find a single or very limited number of broadly effective therapy strategies. Abs are one key component of molecularly targeted therapy and there are currently 3 FDA- and EMA-approved mAbs directed to breast cancer treatment, all regarding HER2-positive tumors (Landes Bioscience, 2014).

Phage display technology plays an important role in the generation and identification of highly specific human mAbs (Hoogenboom et al., 1998), and it has also revealed appealing for breast cancer research. Among the explored applications for this technology in the area are the identification of new targets (Minenkova et al., 2003), epitope mapping (Gabrielli et al., 2013), identification of peptides with affinity for breast cancer cells (Feng et al., 2014; Fu et al., 2014; Lu et al., 2013) and related targets (Dai et al., 2014) for therapeutic and diagnostic purposes, as well as Abs with therapeutic potential (Schier et al., 1995).

Some investigators have constructed phage display libraries specifically for this end. Pavoni et al (2007) constructed several scFv libraries from breast cancer patients' biologic material, namely three deriving from tumor-infiltrated B lymphocytes of different individual donors, one with the same type of cells but from four different donors and one library from the peripheral blood lymphocytes of one donor, and selected them against known tumor antigens. Ayat et al (2013) generated two scFv libraries from lymph nodes of breast cancer patients that expressed HER2 and CEA antigens in their tumors, and selected them against these same antigens. Rothe et al (2004) also constructed a scFv library from lymph nodes of three breast cancer patients, but this was intended for the isolation Ab fragments that recognize potentially new antigens for this disease.

## 4.1 Libraries Construction and Characterization

Although the examples from the literature described above mention mostly scFv libraries, in this work, two immune phage display Fab libraries (κ and λ) were created in order to obtain a tool for the development of new therapeutic Abs for breast cancer. Unlike scFv's, Fab's are well known fragments that do not tend to multimerize, which facilitates Ab selection and characterization, and allows the isolation of high affinity clones. Furthermore, Fab conversion into whole Ab molecules has a less significant impact in the initial molecule functionality (Burton, 2001).

Regarding the proposed application, the Ab gene source chosen for the library repertoire construction was the peripheral blood B cells of 16 breast cancer patients. Evidence has been found that tumor cells elicit a humoral immune response in patients (Pavoni et al., 2006; Pavoni et al., 2007); consequently, breast cancer donors are expected to provide genetic information of highly specific Abs against breast tumor-related proteins for the library construction. On the other hand, considering each individual response peculiarities, breast cancer heterogeneity and molecular variability, using a relatively high number of donors is expected to increase the Ab repertoire variability in terms of different Abs for a single antigen, as well as Abs for a considerable amount of different antigens.

Besides the number of donors, other strategies were used to enhance the constructed repertoire's variability. For HC Fd and LC fragments construction, multiple segments were first separately amplified, with several primer combinations, and then subsequently randomly merged. The use of several primer combinations intended to preserve the original repertoires diversity by allowing the access to as many different V-gene segments available from the recovered cDNA as possible (Burton, 2001; Hammers and Stanley, 2014; Tohidkia et al., 2012). On the other hand, the random assembly of segments by overlap PCR, as well as LC and HC Fd random combination by sequential cloning, should significantly contribute to diversity maximization.

The libraries final characteristics and application success will not only depend on its design but also on the techniques and the outcome of each step of the construction (Shahsavarian et al., 2014; Tohidkia et al., 2012), which starts with the donors' biologic material recovery and RNA preparation. RNA isolation yield varied considerably between samples. However, common side-effects of cancer treatment, including several chemotherapy drugs, radiotherapy and even some targeted therapies are neutropenia and lymphopenia (Breastcancer.org, 2014b; van der Most, 2009). The only information gathered about the donors was their age and sex, therefore it was not known if some were suffering from this type of side-effects. Nevertheless, it is possible that the patients presented considerably variable numbers of mononuclear blood cells, from which RNA was isolated, resulting in RNA yield fluctuation. RNA purity of the samples was also assessed through A260/A280 ratio spectrophotometry measurement, which is generally expected to be close to 2,0. The results we obtained presented values in the range 1.3-1.94. However, the value A260/A280=2 is standardized for RNA eluted in TE buffer (pH 8.0), while the samples we prepared were eluted in nuclease-free water, which may have a lower pH. It has been shown that even small changes in the pH of the solution may cause the ratio to vary, more specifically, acidic solutions will under-represent it by 0,2-0,3 (Biomedical Genomics, 2007; Thermo Fisher Scientific, Inc, 2008) and thus this could explain the slight difference obtained. Furthermore, RNA treatment with DNase TURBO™ DNA-free Kit (Ambion®) may slightly decrease the ratio when the RNA concentration is lower than 25 ng/μL (Life Technologies Corporation, 2012). Indeed, the samples that presented lower ratios were those that were quantified below 25 ng/μL, whose measurement might have been cumulatively decreased by these two reasons. The remaining samples all had ratios above 1,77, hence it was considered that purity levels were acceptable for all cases. Isolated RNA integrity was preserved, as indicated by the presence of 18S and 28S ribosomal RNA species in agarose gel electrophoresis, with the last one being more intense, as expected.

In order to have the same contribution from each donor to the library, the same quantity of each sample's RNA was used to form the pool from which the repertoires were constructed. Although the RNA quantity contained in the pool was known, the quantity of Ig mRNA it contained and to which the specific primers used for cDNA synthesis could anneal was not. Consequently, the produced cDNA quantity, which also depended on the reverse transcription reaction efficiency, was not known either. This was used as template in the first step of amplification of each gene repertoire. As a different specific cDNA was produced for HC Fd fragment, κ LC and λ LC construction, the template

quantities in this step might have been different for each repertoire, influencing the final yield of the reactions.

During the HC Fd and LC fragments construction process, some PCRs rendered lower amounts of products. In these cases, attempts were made to optimize the reactions in order to improve their yields, however with limited success. Nevertheless, variation in the amount of PCR products is not unexpected, once the reactions' results depend on the efficiency intrinsic to each primer combination, and the amount of template matching to that primer combination available from the pool (as explained before) (Hogrefe and Shopes, 1994). Furthermore, V-gene length variations (especially in the HC CDR3) cause the gel electrophoresis bands too look broader than those formed by typical uniform amplification products. Indeed, sharp bands instead of broad bands could be an indication that the complexity of the amplified Ab repertoire is not very high (Welschof et al., 2003).

The Ab fragments amplification and assembly resulted in the construction of three DNA repertoires: HC Fd, κ LC and λ LC. Fab libraries were constructed by cloning these repertoires into pCOMB3XSS, a phagemid vector appropriated for Fab phage display. This vector also facilitates downstream analysis of selected clones, once it allows the expression of sFab's that carry tags for purification ($His_6$-tag) and detection (HA-tag).

Inserts restriction enzyme digestion efficiency is critical for the library final size and diversity, once undigested fragments cannot be ligated to the vector and therefore will not be integrated in the library (Shahsavarian et al., 2014; Tohidkia et al., 2012). Inappropriate digestion will reflect in the number of ligation-transformed bacteria, so small scale trials are not only useful to assess the best vector:insert molar ratio for ligation and the vector self-ligation background levels, but also to check if the inserts were appropriately digested. Therefore, although several transformations can be performed to increase the library size, the transformation efficiency should be higher than $1x10^7$ transformants/μg DNA (Clark, 2002; Marks and Bradbury, 2004b). Transformation efficiency must always be evaluated taking into account the competent cells own transformation efficiency, to allow determining if the possibility of obtaining poor results is not caused by low cell competency, but instead due to problems with DNA digestion, dephosphorylation, and/or ligation. For this purpose, a transformation control, like a pUC plasmid, for example, pUC19, is usually used, and in this test the transformation efficiency in highly competent cells is expected to be higher than $1x10^9$ (Marks and Bradbury, 2004b), to exclude problems concerning the transformation itself. By performing small scale trial ligations, it was possible to confirm that all of these conditions were regarded. Vector self-ligation levels, on the other hand, should be lower than 0,1% (Clark, 2002) to minimize the library portion constituted by phagemid not containing Fab genetic information. This condition was not strictly fulfilled for the Hc Fd fragments' cloning into the vector already containing the LC. However, after a second and more careful gel purification and a longer dephosphorylation reaction of the digested LC-containing vectors, the self-ligation background levels were only marginally higher than 0,1% and therefore considered acceptable.

Immune libraries should comprise at least $1x10^7$ transformants, which is less than what is required for naïve libraries ($1x10^9$) (Marks and Bradbury, 2004b), once these are more dependent on the size to provide high affinity and specific Abs (Bazan et al., 2012; Hoogenboom et al., 1998;

Hoogenboom, 2005). Small scale trials' results also indicate how many transformations are approximately needed to construct a library of the desired size with each particular vector-insert combination. This is usually achieved by transforming all the ligation products available, and thus allowing all potential components in the DNA library to be present in the final phage repertoire (Kotlan and Glassy, 2009). This strategy was followed and the sizes of the newly generated libraries was in the $10^8$ range. Other reports have shown similar values for immune library sizes: for example, the libraries constructed by Pavoni et al (2007) sizes ranged from $4{,}7x10^5$ to $2{,}6x10^7$ clones, Rothe et al (2004) obtained a repertoire of $5x10^8$ clones, and human Fab libraries generated from patients with other malignancies have been reported, comprising $1{,}3x10^6$ to $1{,}45x10^8$ transformants (Clark et al., 1997; Coomber et al., 1999; Coomber and Ward, 2001; Dantas-Barbosa et al., 2005). As such, we can conclude that the sizes of the two newly immune libraries constructed in this work are standard comparing to other human immune libraries derived from cancer patients material. Also, comparing to immune libraries in general, and according to Frenzel et al (2014), the size of our κ and λ immune libraries are in the acceptable range, since the author indicates that these types of libraries usually comprise sizes between $10^6$-$10^8$ independent clones.

After the construction an initial validation step was followed, and the libraries were characterized by assessing the percentage of clones that were estimated to contain a correctly constructed Fab fragment. Besides the input of possible self-ligated vectors (containing no Fab inserts), incorrectly constructed vectors may also contribute significantly to the undesired library components. This occurs mainly because smaller phagemids (for example, due to deletions or cloning errors) are preferentially transformed and later packaged into the phage coat (Scott, 2001). A small but significant sample of each library (100 clones) was analyzed regarding the number of correctly constructed Fab fragments. Our results indicated that the κ library contained 51% correct clones, while the λ library presented 52%. Other authors have reported variable results, like 40% (Wu et al., 2010), 90% (Rothe et al., 2004) or 100% (Dantas-Barbosa et al., 2005). Although some of these numbers appear substantially higher than those obtained in this work, it is important to emphasize that the study from Rothe *et al* (2004) does not indicate the number of clones tested. Moreover, Dantas-Barbosa *et al* (2005) only analyzed a very small sample of 10 clones for Fab presence, which is scarcely representative, and thus the value reported could be easily altered with a higher number of clones tested. The same is verified for diversity, as in both cases, these authors obtained 100% of diversity by DNA fingerprinting of only 10 clones. In this work, *Bst*NI fingerprinting analysis was performed to more than 40 clones of each library, and the data obtained suggested that over 85% of the clones were different, which means the libraries are highly diverse. Diversity was further analyzed by sequencing, which is the ultimate test and provides additional information such as the inserts identification (and thus presence confirmation) and allows determining if they are full-length and free of cloning errors (mutations, deletions, frameshifts, among others). The high diversity of the generated libraries was not only confirmed by the sequencing results but also reinforced (>91% of different clones), once it was verified that some of the clones whose fingerprinting pattern was identical actually presented different sequences. This supports the fact that, comparing to libraries of less restricted diversity (like naïve), fingerprinting analysis is less sensitive when performed in immune repertoires,

whose diversity (partially enhanced by affinity maturation) might not be translated in gain and loss of restriction sites (Bradbury and Marks, 2004).

The correctly sequenced clones were further analyzed recurring to VBASE2 software (Retter et al., 2005) that allows sequence analysis, providing information about the prevalence of possible errors, the position and sequence of key features (like the FRs and CDRs) and the V-genes source. Analyzed clones were related to most of the human V-gene families and a significant number of correspondent V genes. However, it is not known if the detected distribution was representative of the original repertoire, once it was derived from various individual repertoires of donors that should be biased by each unique immune history, besides the common factor of having breast cancer.

The CDR3 regions of both the HC and LC were also analyzed as an indicator of diversity, given its importance to Ab variability, specificity and antigen binding. Of all the CDRs, CDR3 is the most diverse, particularly that of VH region which also contributes the most to antigen binding, since CDR3 residues are responsible for most of the surface contact area and molecular interaction with the antigen (Goldsby et al., 2003; Murphy, 2012). Clones' CDR3 analysis results were in accordance to these facts, once sequences were highly variable and HC CDR3 length distribution (9 to 22 amino acid residues) was clearly wider than that of LC CDR3 (9 to 13 amino acid residues). Human immune libraries described by other authors present similar distributions. Unselected clones from a Fab immune library constructed from the repertoires of ten donors vaccinated against rabies presented HC CDR3 with lengths from 6 to 18 amino acid residues (Houimel, 2014). Sblattero and Bradbury (1999) reported a scFv non-immune library with CDR3 lengths ranging from 7 to 22 amino acid residues for the HC and 8 to 11 for the LC. Another scFv non-immune library presented a HC CDR3 length distribution from 5 to 18 amino acid residues (Sheets et al., 1998).

Besides the importance of library size and diversity, genetic characterization may not reflect the functionality (Carmen and Jermutus, 2002). A clone that contains a complete and in-frame Fab sequence may not result in the production of a functional protein if the expression system is not compatible, for example due to protein interference with the phage assembly, cell toxicity or folding machinery limitations (Carmen and Jermutus, 2002; Hammers and Stanley, 2014). The own characteristics of the phagemid system results in a high frequency of monovalent display, so the vast majority of phage will display only one copy of the Fab, or even none (Carmen and Jermutus, 2002). This is an advantage if the purpose is to select for high affinity binders (Ahmad et al., 2012), but an efficient phagemid rescue is needed to keep the display levels high enough so that the library's functionality is not lost and its potential is still represented (Tohidkia et al., 2012).

After libraries rescue, phages were shown to be present in high quantities (even exceeding the ELISA sensitivity limit in the conditions tested), and their display was as expected since there was expression of Fab's in a dose dependent manner. When comparing the two immune libraries generated, the κ library appeared to have slightly higher Fab expression levels. Although a pronounced difference between the signals obtained with the anti-LC Abs was verified, it has to be taken into account that two different Abs (for each library LC specificity) were used, and their intrinsic characteristics may influence the displayed Fab's detection results. Altogether, these results indicated

that the rescue protocols followed were successful and phage libraries aliquots were prepared to be subsequently used for panning.

## 4.2 Libraries Selection

Selection of Abs or other peptides and proteins with specific binding affinities from millions of variants is the whole purpose of phage display technology (Ahmad et al., 2012). Therefore, a phage display library is only fully validated after testing its performance in phage panning selections (Carmen and Jermutus, 2002). In the present study, selection was performed using three different targets. For general performance validation, libraries were tested against BSA, a foreign, stable, and highly immunogenic protein (Hayworth, 2014). Secondly, HER2, a well-established breast cancer biomarker and Ab therapeutic target, was chosen for the library concept validation. The last selection target was DLL1, given its involvement in the Notch signaling pathway, and due to the fact that the aim of this dissertation is to generate function-blocking Abs against a ligand of the Notch1 receptor.

In a 1$^{st}$ round of panning, a direct library sample, with a determined number of phages ($10^{12}$ in our case) is used. In this initial step all the clones present in the library are first subjected to a depletion step (if available), and the remaining phages are carried over to the selection with the specific antigen. In the subsequent rounds the input phage sample will not present the initial variability of the library, but will be composed of phages that were recovered from the previous round, and that are more specific towards the target. Therefore, it is crucial that the initial sample represents the full library diversity and that each different clone is present with sufficient frequency (Kotlan and Glassy, 2009). As the vast majority of phage is monovalent and a significant number may not be displaying at all, large populations of phage, much greater than the library size, are required for an efficient representation of the library diversity (Carmen and Jermutus, 2002); indeed, Kotlan and Glassy (2009) recommend the addition of sufficient phages to ensure at least 1000-fold representation. However, there are limitations considering the practical number of phages that can be used in selection, namely $1x10^{13}$ phages in 1 mL (Sblattero and Bradbury, 2000). Usually, if there is enough antigen available for panning, the selection steps are carried out in a final standard volume of 1 mL. However, in the current strategies we were faced with scarcity of antigen, and thus we down-scaled the three panning procedures. As such the selection strategies were carried out in 96-well Maxisorp plates in a final volume of 0,1 mL or 0,05 mL, depending on the antigen used.

Although it is always defended by several authors that a highly diverse library must be used for panning, this may imply that there are relatively few copies of each individual displayed Ab (Marks and Bradbury, 2004a). As such, other strategies can be pursued to overcome this, in our case the phages were selected twice for the same antigen in the 1$^{st}$ round, in order to increase the chances of interaction between specific binders and the target.

The number of phages that will be recovered from the 1$^{st}$ round will depend on the quantity of clones that bind to the antigen and on the stringency applied (Kotlan and Glassy, 2009; Marks and Bradbury, 2004a). 1$^{st}$ round results of the performed selections were typical (between $10^4$–$10^6$ recovered phage particles, which is expected to be between $10^4$–$10^7$ according to Marks and Bradbury (2004a). While the output phage titers and percentage of input recovery should increase between rounds of selection, indicating selection is occurring and leading to phage enrichment for the specific

target (Marks and Bradbury, 2004a), in some cases it is also typical that the output phage titers are lower in the 2nd round (Rader, 2012). This is because stringency is usually increased at this stage to favor the removal of low affinity and nonspecific bound Abs. After the 1st round of selection, phage diversity should have considerably diminished and at the same time the number of each individual clone should be far higher due to amplification of the recovered subset. Therefore, it is safe and productive to increase the stringency, so that the output will comprise a higher proportion of positive phages (Marks and Bradbury, 2004a). In this work, stringency was increased by doubling the number of washing steps from 5 in the 1st round to 10 in the following. Consequently, it was observed that 2nd round output phage titers were lower than 1st rounds'.

### 4.2.1 BSA

BSA panning was the first to be performed, and for this strategy we chose to use $1x10^{12}$ phages in 0,1 mL in the 1st round of selection. However, using this number of phages implied that there was an unequal representation of the two libraries, considering their different size – approximately 590-fold representation of κ library (below recommended) *versus* 2180-fold of λ library. Furthermore, after real input calculation it was found that the λ library input value ($1,27x10^{13}$ phages) was approximately 10-fold higher than κ library's, contrarily to the intended, and therefore the representation difference between the two libraries was significantly increased. The unexpected λ library input value was possibly caused by a dilution error.

Regarding the output titer for the λ library, this value increased approximately 190-fold from the 2nd to the 3rd round and a phage enrichment of over 190-fold was also observed, suggesting that the panning was successful. For the κ library selection, the 3rd round output values were acceptable (in the $10^5$ range) but the values for the 2nd round were lower than usual (in the $10^2$ range), which caused a difference of about 1000-fold between the two rounds and an enrichment value atypically high (over 17000-fold). Although, as referred, output values can decrease from the 1st to the 2nd round, in this case it was not inside the usual parameters, but considering that the 3rd round values were in the expected range, it was assumed an error might have occurred in the 2nd round output tittering. Therefore, percentage of recovery of the 2nd round and 3rd round enrichment value should not be considered. In order to clarify this value the titration of the output phages from the 2nd should be repeated.

Selected phages from all rounds were tested to verify if target-specific Fab's were being recovered for both libraries. The analysis of amplified BSA-selected phages from the κ library indicated that the selection failed, and although Fab presence was detected, it seemed to decrease from round to round and BSA reactivity was not shown in any case. Some phages might have been recovered in the 1st round, once the phage recovery was higher than that of λ library for the same round. However, there might have been loss of some clones in the elution step, if the TEA incubation time was not the ideal. While phages may not be recovered if this incubation is too short, they may be degraded if it takes too long (Charlton and Porter, 2002). Another hypothesis is that the phages recovered from 1st round were not specific for BSA, and the higher stringency in the following rounds led to their progressive loss. Any of these situations could explain the 2nd round unexpected output value and low phage recovery, initially thought to have been caused by a tittering error. Considering

the second hypothesis, a possible reason for the panning not having succeeded is the underrepresentation of the library diversity on the 1st round phage input. This means that better results could be achieved by starting the selection with a higher number of phages despite the practical recommendations above mentioned, or by applying a different strategy that would allow the use of higher volumes. On the other hand, the κ library itself may not be as functional as the λ; quality and characterization tests performed do not indicate that, once both libraries had similar diversity and Fab displaying results; the noticeable exception is that κ library, despite being bigger, may contain a higher percentage of Fab's containing errors than λ.

The ELISA analysis of λ library amplified phages pools results suggested that there was an enrichment of BSA-specific phages from the 1st to the 3rd round, as the pools reactivity for the target progressively increased, contrarily to the verified for κ library. Considering these results, individual clones selected from the λ library were tested for specificity. Clones that were considered positive in Fab-on-phage ELISA showed signals for Fab presence and BSA-reactivity far stronger than the remaining tested clones. Positive signals were maintained on sFab ELISA but in more modest levels, probably due to this test being dependent on the expression yield of the Fab fragments in soluble format, which might not be optimized.

Further tests were performed to characterize the positive clones. Genetic studies aim to assess if binders with affinity for a certain target are different or not or if they share features that may determine their affinity for a certain target. Sequencing is the more reliable test, once it reveals the Fab's coding information and therefore allows an absolute comparison between clones. However, before sequencing, BSA-selected clones were amplified by PCR and the results analysis by gel electrophoresis showed a band with a molecular weight 300 bp-400 bp higher than expected. Furthermore, *Sac*I/*Spe*I digestion of the phagemid DNA also presented bands of unexpected lengths, once the vector moiety should have the same size as the correspondent fragment resulting from the digestion of pCOMB3XSS vector, which was used as a control, but was approximately 1 kb bigger instead. These unexpected results suggested that the phagemids were not correctly constructed, and due to the patterns obtained we could even suggest all clones were the same. This might have been caused by overrepresentation of an incorrectly constructed clone in the selection output, which may have superior growth kinetics, resulting in its dominance after several rounds of enrichment of an increasingly restricted population. Although performing additional rounds of selection should increase the number of positive clones, in this case it could enhance the dominance of the defective clone. Therefore, the panning should be repeated and other strategies tested in order to obtain diverse output populations from the beginning, when the loss of strong candidates is critical (Carmen and Jermutus, 2002).

### 4.2.2 HER2

Considering the results obtained for the κ library selection against BSA, only λ library was subjected to panning against HER2 and DLL1, due to limitations in time and resources.

For HER2 and DLL1 pannings, $1x10^{12}$ of phages in a 0,05 mL volume were theoretically used in the 1st round. Due to limitations in terms of the target quantity available, it was decided to perform the pannings in a smaller volume than that of the BSA selection (0,1 mL). However, in order to keep

the library representation levels, the number of phages was not reduced. Although it implies that the recommendation in terms of phage concentration was not strictly fulfilled, the limit was not substantially surpassed, once $2x10^{13}$/mL were used instead of $1x10^{13}$/mL.

In the HER2 selection, the output titer increased approximately 190-fold from the 2nd to the 3rd round, which translated in a phage enrichment of 364-fold, suggesting panning success. Subsequent analysis of the HER2-selected phage pools from the 2nd and 3rd round showed specific dose responsive HER2 reactivity. This suggested that target-specific Fab's were recovered, although the Fab display results using the anti-Fab specific Ab were not typical in terms of dose-response. 3rd round HER2-selected clones were tested for specificity and, identically to what was verified for BSA, several clones were found positive both on Fab-on-phage and sFab ELISAs, showing stronger signals for Fab presence and HER2-reactivity in the first test.

Moreover, gel electrophoresis analysis of HER2 clones PCR products showed no bands for H9 clone and a band of approximately 1700 bp for the remaining clones, instead of the expected 1500 bp. *Sac*I/*Spe*I digestion pattern was not the expected either. This suggested that, identically to what was verified for BSA, all HER2 positive clones tested presented construction errors. It is possible that this selection strategy might also favor defective clones, which compose about 50% of the libraries, corroborating the need of trying different panning approaches. For HER2 selection, 2nd round polyclonal ELISA showed results in a certain way better than the 3rd round's. Therefore, testing clones from the 2nd round output could result in finding positive specimens with more interesting characteristics. Nevertheless, the 3rd round or the whole panning should be repeated, and other strategies tested, as referred.

These preliminary conclusions about BSA and HER2 positive clones should be in any case confirmed by sequencing. Therefore, clones' phagemid DNA was sent for sequencing, but the results were not satisfactory for both BSA- and HER2-clones. This DNA presented good values in terms of quantification and purity (1,78-1,86) when assessed by spectrophotometry, and when ran in an agarose gel electrophoresis, several bands correspondent to the phagemids in different conformations were observed as expected. Since these results suggested that there were no obvious problems with the DNA, the samples were sent for sequencing in a different company, but once more the results quality did not allow an analysis. In a third attempt, sequencing was tried using as template Fab insert amplification products, but it was not successful either. Despite the mentioned difficulties, more alternatives should be tried before giving it up, namely repeating the phagemids isolation and retrying the sequencing with the new DNA batches. Before sending the DNA to sequence, its concentration could be assessed by comparison of the intensity of the bands resulting from running it in an agarose gel with that of the bands of a DNA sample of known concentration, once some sequencing companies consider this method is more feasible than spectrophotometry quantification.

### 4.2.3 DLL1

The recent association between deregulation of Notch signaling and breast cancer (Joo et al., 2013; Mittal et al., 2009; Reedijk et al., 2005; Rizzo et al., 2008; Sharma et al., 2012) has led to a great interest in unraveling the various mechanisms of this association (Izrailit and Reedijk, 2012; Joo et al., 2013; Sethi et al., 2011; Takebe et al., 2014), and exploring the potential of the pathway as a

therapeutic target (Sharma et al., 2012; Simmons et al., 2012; Suman et al., 2013; Wang et al., 2011). As a result, different strategies are being developed to block Notch signaling for therapeutic purposes, and several Notch-targeting mAbs are currently in early phase clinical development (Takebe et al., 2014). A great part of the research on this matter has been centered on the pathway's receptors and the first steps in understanding the specific role of some ligands, like Jag1 and DLL4, have already been taken (Izrailit and Reedijk, 2012; Sethi et al., 2011). On the other hand, little is known about DLL1 ligand's implication in the breast cancer malignancy condition, although evidence on its existence has been found. Several authors have studied patients' samples in order to evaluate the presence of DLL1. Reedjik *et al* (2005) identified DLL1 mRNA expression in 9,1% of the analyzed human breast cancers cases, and Rizzo *et al* detected high DLL1 expression in 22,2% and 36,7% of ductal infiltrating carcinomas and lobular infiltrating carcinomas samples, respectively. Approximately 81% of invasive ductal carcinoma cases examined by Mittal *et al* (2009) showed low to moderate expression of DLL1, while they failed to detect any significant expression in normal breast tissue. High expression of DLL1 was also found in 39,0% of cases analyzed by Yanai et al. (2010). The isolation of Abs against DLL1 ligand would constitute one more tool to study the relation between the Notch pathway and breast cancer by improving the knowledge underlying the mechanism of action, and evaluating the therapy target potential of a promising ligand that has been little explored to date.

In this work a portion of the extracellular domain of the DLL1 protein (Trp 159 – Val 444, domains DSL to EGF-like 6) fused to a $His_8$-tag on its N-terminal was used as a target for selection of Abs by phage display, using the newly generated λ library. In a parallel work, several DLL1 protein constructions were produced in several vectors, and small scale protein expression tests were carried out both in *E. coli* and mammalian cells in order to assess and optimize expression yields. The DLL1 protein used for the panning was produced from HEK293-EBNA6 cells since this system presented slightly higher expression levels. This work was developed in the context of my colleague Margarida Silva Master Thesis and will soon be discussed elsewhere.

DLL1 selection showed an output titer increase of approximately 190-fold from the 2$^{nd}$ to the 3$^{rd}$ round, which translated in a phage enrichment of 195-fold, suggesting panning success. DLL1-selected phage pool analysis results were inconclusive; while 2$^{nd}$ round outcome was satisfactory, the lack of dose response in the signal from 3$^{rd}$ round phages test was not. However, there was still evidence of DLL1 specific reactivity. Since the phage recovery and enrichment values calculated suggested that an efficient selection occurred in the 3$^{rd}$ round and the ELISA results did not excluded the presence of antigen-specific Abs, individual clones were tested for specificity to confirm the data. None of the clones tested against DLL1 clearly highlighted from the remaining, but 3 had stronger signals for DLL1 reactivity on Fab-on-phage ELISA (at least 2,3-fold over the background). Similarly to what was observed in BSA- and HER2-selected clones sFab test, DLL1-selected sFab signal was weaker than that of Fab-on-phage ELISA. Due to this decrease, DLL1-selected sFab clones signal was very weak and only slightly higher than background. Nevertheless, clones' characterization tests were also performed for these clones, hoping to find a binder, which even with a present low affinity could be affinity maturated for improvement. Affinity maturation is frequently used to optimize Abs initially selected for a target by introducing random or directed mutations in the clones' V-genes and

subsequently selecting the variants of higher affinity by a process similar to panning. This usually results in Ab fragments with affinities in the sub-nanomolar range (Hoogenboom et al., 1998; Schofield et al., 2014).

The genetic analysis of DLL1-selected clones was more elucidatory than that performed for the clones obtained in the other selections, particularly because sequencing information was obtained. D2 and D19 clones were correctly constructed and shared the same sequence, which could have been predicted by the *Bst*NI digest. However, these clones have not shown a very strong reactivity for DLL1, and Fab expression was also barely detected by ELISA and Western Blot. Although the Fab sequence does not contain errors, it is possible that this clone is expressed at low levels, for example due to sensitivity to bacterial proteases (Tohidkia et al., 2012), toxicity of the clone to the host cell or interference with phage assembly (Hammers and Stanley, 2014). Therefore, even if they have affinity for DLL1, they may not be present in quantity enough for being detected in the assays. On the other hand, D3 clone's Fab insert was not correctly constructed, which was predicted, since the pattern detected on agarose gel electrophoresis for the whole phagemid was discrepant from the expected (comparing to D2/D19 and pCOMB3XSS pattern, which is similar to that of correctly constructed clones). Also, the *Sac*I/*Spe*I digestion products presented a different size too, indicating that there was no Fab fragment present in this clone but a deletion mutant that explains the smaller molecular weight. Sequencing was only obtained with the forward primer, and the results obtained agreed with the remaining tests by showing that the LC insert started on CDR1, so FR1 was absent. LC inserts were cloned by digestion with *Sac*I and *Xba*I enzymes. Although the *Xba*I site was found, it was not followed by the expected sequence that exists between *Xba*I site and *Xho*I site and includes the *pelB* leader. On the other hand, 133 bp downstream of the *Xba*I site a different part of the pCOMB3XSS sequence was detected, so it is possible that part of the phagemid sequence was deleted, namely the portion that includes the HC insert and Rv primer site. Therefore, only the LC was being expressed and possibly had some affinity for DLL1.

Antigens titration ELISA showed BSA and HER2 proteins were in good condition and therefore should not hamper the panning success. Conversely, DLL1 failed to be detected with a specific primary mAb. This mAb (Sino Biological Inc., Cat# 11635-MM07) was produced by hybridoma technology; B cells were obtained from a mouse immunized with recombinant human DLL1 extracellular domain (Met 1-Gly 540). The protein used in this work, on the other hand, is only part of the extracellular domain (Trp 159 – Val 444), so it is possible that the mAb recognizes an epitope that is not represented in that specific DLL1 portion. The protein detection was also performed with a primary anti-His Ab and a slight, increasing signal was detected for the highest dilutions. Also, a single aliquot of protein was used for these procedures and it was always kept at 4ºC to avoid consecutive freezing and thawing cycles. As such, the current batch may have lost quality along time, thus explaining the inconsistency between the 2$^{nd}$ and 3$^{rd}$ round results. Protein instability is not surprising once this DLL1 is a highly complex protein with a particular pencil-shape conformation bearing innumerous pairs of disulfide bonds (Uniprot, 2014). This protein construction was produced in house and its expression and purification presented some challenges. Nevertheless, better results could be

achieved by repeating the 3rd round or the whole panning with another batch of DLL1 after testing the protein by ELISA (including detection with a Anti-his Ab to prevent false negatives), so that its quality is verified prior to panning.

The chemistry of the antigen immobilization may also be responsible for the DLL1 protein not maintaining its native structure, since antigen is usually coated onto solid phase by passive adsorption, which may cause denaturation of the molecules. Proteins denatured forms may expose hydrophobic residues, causing a 'sticky' surface that in some cases can favor non-specific phage or antibody binding (Carmen and Jermutus, 2002). The density of immobilized antigen may also influence epitopes availability (Carmen and Jermutus, 2002), so the concentration used for coating with this target may not have been adequate. Performing pannings with various DLL1 target concentrations may be needed to find the ideal conditions.

Although first panning results indicated it was proceeding successfully, isolation of Fab clones with characteristics of interest was not accomplished for any of the selections performed, once initial results indicate that the BSA and HER2 clones that showed signs of specific binding contain sequence errors and the DLL1 tested clones are expressed at low levels. Input recovery and phage enrichment values are not absolute indicators, once cases of successful selection with no change in output titer and failed selections with increases in titers have both been observed (Marks and Bradbury, 2004a). In any case, polyclonal ELISA tests also corroborated target-specific phages were being isolated and it is important to notice that only a small number of individual clones was tested for each selection, so it was not widely representative; thus, by testing a more numerically relevant sample, different positive clones could have been found. Altogether, it can be foreseen that by repeating the selections after optimization of the conditions, including the performance of preliminary tests to assess the optimal of input phage quantity and target concentration, more satisfactory results will be achieved and ultimately result in the isolation of Ab fragments of interest for further research.

## 4.3 Conclusions

In summary, the work reported herein resulted in the construction of two immune phage display Fab libraries (κ and λ) containing more than $10^8$ independent clones and a diversity of over 90%, characteristics equivalent to those of the immune libraries reported by other studies. The two first goals of this study were, therefore, achieved. Regarding the libraries validation and selection of specific Fab's against the DLL1 protein, only preliminary results were obtained, so further effort will have to be put in this tasks in the future. Nevertheless, a tool for the isolation of Ab fragments specific for breast cancer-related targets was successfully created; given its countless applications, including the development of new therapeutic agents, it constitutes a positive contribution to research in this area.

## 4.4 Future perspectives

With the work developed in this thesis we were able to successfully generate two immune Ab human Fab libraries (κ and λ) with a high diversity. However, a bigger investment must be still carried out regarding the validation of both libraries upon panning selection. Due to the time available for the development of this study, a complete validation process was not carried out as thorough and as

successfully as initially planned. The preliminary data obtained with the BSA panning did not render the expected results, but we still decided to proceed simultaneously with the pannings using HER2 and DLL1 antigens, hoping to enrich this dissertation.

The next steps essential for the continuation of this work will be first to understand why the BSA and HER2 sequencing failed. For this task new DNA preparations could be carried out, or even the selection of more individual clones could be pursued. Obtaining the sequence for these clones will allow clarifying the questions that arose from the previous genetic tests.

Besides that, the next critical step would be to repeat the libraries selection employing other panning strategies. Different concentrations of antigens, or even different antigens could be tested, and the selection procedure could also be altered for example to a selection with beads.

Regarding the clone we obtained for the DLL1 panning the following work will be to further characterize it. Also, other clones can be picked and assessed for DLL1 binding, or even another selection can be carried out to increase the number of lead candidates. After isolating DLL1-specific Fab's with the desired characteristics, these will be expressed in a relatively larger scale and purified in order to be further tested. Selected clones affinity will be measured using the Biacore or the Octet, and their performances compared to allow a ranking of the Fab's. Ab-antigen interaction will also be characterized, in terms of determining the specific epitopes and unraveling Ab-antigen complexes tri-dimensional structure. The Fab clones with the best characteristics will be reformated into the complete IgG molecules, and those will be subjected to identical studies so that possible behavior alterations can be identified. Potential therapeutic agent candidates will be screened for biological function in cell-based bioassays.

Since this work is integrated in a broader project, in parallel to this work a Fab naïve library has also been constructed for the same purposes as the immune library reported herein. Therefore, the two libraries performance in selection and the Ab fragments isolated from both will be compared in order to investigate to what extent the type of library can influence the outcome of an Ab selection against tumor-specific targets. As one of the aims of the referred project is to characterize Notch1 ligands and their interaction with specific Abs, the constructed immune library can also be selected against different constructs of the DLL1 protein and other ligands, like for example Jag1 or Jag2. Studies performed with these Abs will help to understand the mechanisms of interaction between the receptor and the ligands and would increase the possibilities of finding function-blocking Abs.

Although immune libraries application is by principle somewhat limited, the design of the library reported here keeps a wide range of use possibilities, as it was not restrained to a single antigen but instead thought to be used against any breast cancer associated target, thus being a valuable tool for the research in this matter.

# 5. References


Ahmad, Z., et al. 2012. scFv Ab: Principles and Clinical Application. Clin Dev Immunol, 2012: 980250.

Akram, M. and Siddiqui, S.A. 2012. Breast cancer management: past, present and evolving. Indian Journal of Cancer, 49(3): 277-82.

Andris-Widhopf, J., et al. 2011. Generation of Human Fab Antibody Libraries: PCR Amplification and Assembly of Light- and Heavy-Chain Coding Sequences. 2011(9): pii: pdb.prot065565.

Arbabi-Ghahroudi, M., Tanha, J., and MacKenzie, R. 2009. Isolation of Monoclonal Antibody Fragments from Phage Display Libraries. Methods in Molecular Biology, 502:341-364.

Aste-Amézaga, et al. 2010. Characterization of Notch1 Antibodies That Inhibit Signaling of Both Normal and Mutated Notch1 Receptors. PLoS One, 5(2): e9094.

Ayat, H., et al. 2013. Isolation of scFv antibody fragments against HER2 and CEA tumor antigens from combinatorial antibody libraries derived from cancer patients. Biologicals, 41(6): 345-54.

Ballantyne, A., Dhillon, S. 2013. Trastuzumab Emtansine: First Global Approval. Drugs, 73(7): 755-65.

Bazan, J., Całkosiński, I. and Gamian, A. 2012. Phage display – A powerful technique for immunotherapy. Human Vaccines and Immunotherapeutics, 8(12): 1817-28.

Biomedical Genomics. 2007. RNA Quality Control . Version 16 September 2014. http://biomedical genomics.org/RNA_quality_control.html in Biomedical Genomics, http://biomedicalgenomics.org.

Bolós, V., et al. 2013. Notch activation stimulates migration of breast cancer cells and promotes tumor growth. Breast Cancer Research, 15(4): R54.

Biotem. 2011. Explore the Ultimate Humanization™ Platform. Version 13 September 2014. http://www.biotem.fr/anticorps-ultimate-humanization/p58.html in Biotem, http://www.biotem.fr.

Bradbury, A. and Marks, J. 2004. Phage Antibody Libraries. *In* Phage Display (T. Clackson and H.B. Lowman eds) 1st ed., pp.243-86, Oxford University Press, New York.

Bray, F., Ren, J.S., Masuyer, E. and Ferlay, J. 2013. Global estimates of cancer prevalence for 27 sites in the adult population in 2008. International Journal of Cancer, 132(5): 1133-45.

Breastcancer.org. 2014a. Targeted Therapies. Version 27 July 2014. http://www.breastcancer.org/ treatment/targeted_therapies in Breastcancer.org, http://www.breastcancer.org.

Breastcancer.org. 2014b. Low White Blood Cell Count. Version 16-09-14. http://www.breastcancer. org/treatment/side_effects/low_white_blood_cell in Breastcancer.org, http://www.breastcancer.org.

Brissette, R., Prendergast, J.K. and Goldstein, N.I. 2006. Identification of cancer targets and therapeutics using phage display. Curr Opin Drug Discov Devel, 9(3): 363-69.

Burton, D.R. 2001. Antibody Libraries. *In* Phage Display – A Laboratory Manual (C. F. Barbas III, et al eds), 1st ed., pp 11.1-11.24, Cold Spring Harbor Laboratory Press, New York.

Calaf, G., et al. 2014. Vimentin and Notch as biomarkers for breast cancer progression. Oncol Lett, 7(3): 721-727.

Cancer Research UK. 2014a. Types of treatment for breast cancer. Version 27 July 2014. http://www.cancerresearchuk.org/about-cancer/type/breast-cancer/treatment/which-treatment-for-breast-cancer in Cancer Research UK, http://www.cancerresearchuk.org.

Cancer Research UK. 2014b. Biological therapy for secondary breast cancer. Version 1 September 2014. http://www.cancerresearchuk.org/about-cancer/type/breast-cancer/secondary/treatment/biological-therapy-for-secondary-breast-cancer in Cancer Research UK, http://www.cancerresearchuk.org.

Carmen, S. and Jermutus, L. 2002. Concepts in antibody phage display. Briefings in Functional Genomics and Proteomics , 1(2): 189-203.

Charlton, K.A. and Porter, A.J. 2002. Isolation of anti-hapten specific antibody fragments from combinatorial libraries. Methods in Molecular Biology, 178: 159-71.

Clark, M.A. 2002. Standard Protocols for Construction of Fab Libraries. Methods Mol Biol, 178: 39-58.

Clark, M., et al.1997. Isolation of human anti-c-erbB-2 Fabs from a lymph node-derived phage display library. Clinical and Experimental Immunology, 109(1): 166-74.

Coleman, M., et al. 2008. Cancer survival in five continents: a worldwide population-based study (CONCORD). The Lancet Oncology, 9(8): 730-56.

Coomber, D.W., et al. 1999. Generation of anti-p53 Fab fragments from individuals with colorectal cancer using phage display. The journal of Immunology, 163(4): 2276-83.

Coomber D, Ward R. 2001. Isolation of human Abs against the central DNA binding domain of p53 from an individual with colorectal cancer using Ab phage display. Clin Cancer Res, 7(9): 2802-208.

Dübel, S. 2007. Recombinant therapeutic antibodies. Appl Microbiol Biotechnol, 74(4): 723–729.

Dai, X., et al. 2014. Identification of a novel aFGF-binding peptide with anti-tumor effect on breast cancer from phage display library. Biochem Biophys Res Commun, 445(4): 795-801.

Dantas-Barbosa, C., et al. 2005. Construction of a human Fab phage display library from antibody repertoires of osteosarcoma patients.. Genetics and Molecular Research, 4: 126-40.

Dantas-Barbosa, C., et al. 2012. Antibody Phage Display Libraries: Contributions to Oncology. International Journal of Molecular Sciences, 13(5):.5420-40.

de Haard, H. 2002. Construction of Large Naïve Fab Libraries. Methods Mol Biol, 178: 87-100.

de Haard, H., et al. 1999. A Large Non-immunized Human Fab Fragment Phage Library that Permits Rapid Isolation and Kinetic Analysis of High Affinity Antibodies. J Biol Chem, 274(26): 18218-30.

Ebrahimizadeh, W. and Rajabibazl, M. 2014. Bacteriophage Vehicles for Phage Display: Biology, Mechanism, and Application. Current Microbiology, 69(2): 109-20.

Elgert, K. 2009. Immunology: Understanding The Immune System. 2nd ed. Wiley-Blackwell, Hoboken.

Elloumi, J., Jellali, K., Jemel, I. and Aifa, S. 2012. Monoclonal Antibodies as Cancer Therapeutics. Recent Patents on Biotechnology, 6(1): 45-56.

Ersvaer, E.,. 2011. Future Perspectives: Therapeutic Targeting of Notch Signalling May Become a Strategy in Patients Receiving Stem Cell Transplantation for Hematologic Malignancies. Bone Marrow Research, 2011: 570796.

Feng, G., et al. 2014. SPECT and Near-Infrared Fluorescence Imaging of Breast Cancer with a Neuropilin-1-Targeting Peptide. Journal of Controlled Release, 192: 236-42.

Ferlay, J., et al. 2013. Breast Cancer Estimated Incidence, Mortality and Prevalence Worldwide in 2012. Version 9-9-14.http://globocan.iarc.fr/Pages/fact_sheets_cancer.aspx in GLOBOCAN 2012 v1.0, Cancer Incidence and Mortality Worldwide: IARC CancerBase No. 11, http://globocan.iarc.fr.

Frenzel, A., et al. 2014. Construction of human antibody gene libraries and selection of antibodies by phage display. Methods in Molecular Biology, 1060: 215-43.
Fu, B., et al. 2014. Identification and characterization of a novel phage display-derived peptide with affinity for human brain metastatic breast cancer. Biotechnology Letters, Epub ahead of print.
Gabrielli, F., et al. 2013. Identification of Relevant Conformational Epitopes on the HER2 Oncoprotein by Using Large Fragment Phage Display (LFPD). PLos One, 8(3): 58358.
Goldsby, R.A., et al. 2003. Immunology, 5th ed, W.H. Freeman and Company, New York.
Groth, C. and Fortini, M.E. 2012. Therapeutic approaches to modulating Notch signaling: Current challenges and future prospects. Seminars in Cell and Developmental Biology, 23(4): 465-72.
Hammers C.Stanley J.2014.Ab Phage Display:Technique&Applications.J Invest Dermatol,134(2): e17.
Han, J., et al. 2011. Notch signaling as a therapeutic target for breast cancer treatment? Breast Cancer Research, 13(3): 210-17.
Hayworth, D. 2014. Antibody Production. Version 10-9-14. http://www.piercenet.com/method/antibody-production-immunogen-preparation, Thermo Scientific Pierce Protein Biology Products.
Hogrefe, H.H. and Shopes, B. 1994. Construction of phagemid display libraries with PCR-amplified immunoglobulin sequences. Genome Research, 4(2): S109-22.
Hoogenboom H.2005.Selecting and screening recombinant Ab libraries.Nat Biotechnol,23(9):1105-16.
Hoogenboom, H., de Bruïne, A., Hufton, S., Hoet, R., Arends, J., and Roovers, R. 1998. Antibody phage display technology and its applications. Immunotechnology, 4(1): 1-20.
Houimel, M. 2014. The analysis of VH and VL genes repertoires of Fab library built from peripheral B cells of human rabies virus vaccinated donors. Human Immunology, 75(8): 745-55.
Huang, J.X.,. 2012. Development of anti-infectives using phage display: biological agents against bacteria, viruses, and parasites. Antimicrobial Agents and Chemotherapy, 56(9): 4569-82.
Huang, M., Shen, A., Ding, J. and Geng, M. 2014. Molecularly targeted cancer therapy: some lessons from the past decade. Trends in Pharmacological Sciences, 35(1): 41-50.
Izrailit, J. and Reedijk, M. 2012. Developmental pathways in breast cancer and breast tumor-initiating cells: Therapeutic implications. Cancer Letters, 317(2): 115-26.
Jackson, S. and Chester, J. 2014. Personalised cancer medicine. Int J Cancer. Published online in 2-5-14 ahead of print. http://onlinelibrary.wiley.com/doi/10.1002/ijc.28940/pdf in Wiley Online Library.
Joo, W.D., Visintin, I. and Mor, G., 2013. Targeted cancer therapy – Are the days of systemic chemotherapy numbered? Maturitas, 76 (4): 308-14.
Kotlan, B. and Glassy, M.C. 2009. Antibody Phage Display: Overview of a Powerful Technology. Methods in Molecular Biology, 562: 1-16.
Lee, C.M., Iorno, N., Sierro, F. and Christ, D. 2007. Selection of human antibody fragments by phage display. Nature Protocols, 2(11): 3001-08.
Landes Bioscience. 2014. About mAbs Journal. Version 5-9-14. http://www.landesbioscience.com/journals/mabs/about/ in Landes Bioscience, http://www.landesbioscience.com
Life Technologies Corporation. 2012. TURBO DNA-freeTM Kit User Guide. Publication Number 1907M. Revision G.

Lu, R., et al. 2013. Targeted drug delivery systems mediated by a novel Peptide in breast cancer therapy and imaging. PLoS One, 8(6): p.e66128.
Marks, J.D. and Bradbury, D. 2004a. Selection of Human Antibodies from Phage Display. Methods in Molecular Biology, 248: 161-76.
Marks, J.D. and Bradbury, A. 2004b. PCR Cloning of Human Immunoglobulin Genes. Methods in Molecular Biology, 248: 117-34.
Marks, J. D., Hoogenboom, H., Bonnert, T., John McCafferty, J., Griffiths, A., and Winter, G. 1991. By-passing Immunization: Human Antibodies from V-gene Libraries Displayed on Phage. Journal of Molecular Biology, 222(3): 581-97.
Minenkova, O., et al. 2003. Identification of tumor-associated antigens by screening phage-displayed human cDNA libraries with sera from tumor patients. International Journal of Cancer, 106(4): 534-44.
Mittal, S., Subramanyam, D., Dey, D., Kumar, R., and Rangarajan, A. 2009. Cooperation of Notch and Ras/MAPK signaling pathways in human breast carcinogenesis. Molecular Cancer, 23(8): 128-39.
Mukai, H. 2010. Targeted Therapy in Breast Cancer: Current Status and Future Directions. Japanese Journal of Clinical Oncology, 40(8): 711-16.
Murphy, K. 2012. Janeway's Immunobiology. 8th ed. Garland Science, New York.
National Cancer Institute. 2009. Adjuvant and Neoadjuvant Therapy for Breast Cancer. Version 27-7-14.http://www.cancer.gov/cancertopics/factsheet/Therapy/adjuvant-breast in National Cancer Institute.
National Cancer Institute. 2014. Treatment Option Overview. Version 27 July 2014. http://www.cancer.gov/cancertopics/pdq/treatment/breast/Patient/page5 in NCI, http://www.cancer.gov.
Nelson, A.L., Dhimolea, E. and Reichert, J.M. 2010. Development trends for human monoclonal antibody therapeutics. Nature Reviews Drug Discovery, 9(10): 767-74.
Novartis. 2014. AFINITOR® (everolimus) Highlights of Prescribing Information. Version 1 September 2014. http://www.pharma.us.novartis.com/product/pi/pdf/afinitor.pdf in Novartis Pharmaceuticals.
Paul, E.W. 2013. Fundamental Immunology. 7th ed. Lippincott, Williams and Wilkins, Philadelphia.
Pavoni, E., et al. 2006. A study of the humoral immune response of breast cancer patients to a panel of human tumor antigens identified by phage display. Cancer Detection and Prevention, 30(3): 248-56.
Pavoni, E., et al. 2007. Tumor-infiltrating B lymphocytes as an efficient source of highly specific immunoglobulins recognizing tumor cells. BMC Biotechnology, 7(70).
Pazdur, R. 2013. FDA Approval for Everolimus. Version 1 September 2014. http://www.cancer.gov/cancertopics/druginfo/fda-everolimus, in Nacional Cancer Institute, http://www.cancer.gov. USA.
Qi, H., Lu, H., Qiu, H., Petrenko, V., and Liu, A. 2012. Phagemid Vectors for Phage Display: Properties, Characteristics and Construction. Journal of Molecular Biology, 417(3): 129-43.
Rader, C. 2012. Selection of Human Fab Libraries by Phage Display. Methods Mol Biol, 901: 81-99.
Rader, C. and Barbas III, C.F. 1997. Phage display of combinatorial antibody libraries. Current Opinion in Biotechnology, 8(4): 503-08.
Rader, C., et al. 2001. Selection from Antibody Libraries. *In* Phage Display – A Laboratory Manual (C. F. Barbas III, et al eds), 1st ed., pp 10.1-10.20, Cold Spring Harbor Laboratory Press, New York.
Rakonjac, J., et al. 2011. Filamentous Bacteriophage: Biology, Phage Display and Nanotechnology Applications. Current Issues in Molecular Biology, 13(2): 51-76.

Reedijk, M., et al. 2005. High-level Coexpression of JAG1 and NOTCH1 Is Observed in Human Breast Cancer and Is Associated with Poor Overall Survival. Cancer Research, 65: 8530-37.

Retter, I., Althaus, H.H., Münch, R. and Müller, W., 2005. VBASE2, an integrative V gene database. Nucleic Acids Research, 33(Database issue): D671-674.

Ribatti, D. 2014. From the discovery of monoclonal antibodies to their therapeutic application: An historical reappraisal. Immunology Letters, 161: 96–99.

Rizzo, P., et al. 2008. Cross-talk between Notch and the Estrogen Receptor in Breast Cancer Suggests Novel Therapeutic Approaches. Cancer Research, 68(13): 5226-35.

Rothe, A., et al. 2004. Construction of phage display libraries from reactive lymph nodes of breast carcinoma patients and selection for specifically binding human single chain Fv on cell lines. International Journal of Molecular Medicine, 14(4): 729-35.

Sblattero, D. and Bradbury, A. 2000. Exploiting recombination in single bacteria to make large phage antibody libraries. Nature Biotechnology, 18(1): 75-80.

Schier, R., et al. 1995. In vitro and in vivo characterization of a human anti-c-erbB-2 single-chain Fv isolated from a filamentous phage antibody library. Immunotechnology, 1(1): 73-81.

Schirrmann, T., Meyer, T., Schütte, M., Frenzel, A., and Hust, M. 2011. Phage Display for the Generation of Antibodies for Proteome Research, Diagnostics and Therapy. Molecules, 16(1): 412-26.

Schofield, D.J., Lewis, A.R. and Austin, M.J., 2014. Genetic methods of antibody generation and their use in immunohistochemistry. Methods, S1046-2023(14): 00086-3.

Scott, J.K. and Barbas III, C.F. 2001. Phage-display Vectors. *In* Phage Display – A Laboratory Manual (C. F. Barbas III, et al eds), 1st ed., pp 2.1-2.19, Cold Spring Harbor Laboratory Press, New York

Sethi, N., Dai, X., Winter, C.G. and Kang, Y. 2011. Tumor-Derived Jagged1 Promotes Osteolytic Bone Metastasis of Breast Cancer by Engaging Notch Signaling in Bone Cells. Cancer Cell, 19(2): 192-205.

Shahsavarian, M., et al. 2014. Exploitation of rolling circle amplification for the construction of large phage-display antibody libraries. Journal of Immunological Methods , 407: 26-34.

Sharma, A., Paranjape, A.N., Rangarajan, A. and Dighe, R.R. 2012. A Monoclonal Antibody against Human Notch1 Ligand−Binding Domain Depletes Subpopulation of Putative Breast Cancer Stem-like Cells. Molecular Cancer Therapeutics, 11(1): 77-86.

Sheets, M., et al. 1998. Efficient construction of a large nonimmune phage antibody library: the production of high-affinity human single-chain antibodies to protein antigens. Proceedings of the National Academy of Sciences of the United States of America, 95(11): 6157-62.

Shen, Y., Yang, X., Dong, N., Xie, X., Bai, X., and Shi, Y. 2007. Generation and selection of immunized Fab phage display library against human B cell lymphoma. Cell Research, 17(7): 650-60.

Shukra, A.M., Sridevi, N.V., Chandran, D. and Maithal, K. 2014. Production of recombinant antibodies using bacteriophages. European Journal of Microbiology and Immunology, 4(2): 91-98.

Sigma-Aldrich. 2014. Guide to E. coli Markers. V.27-8-14. http://www.sigmaaldrich.com/life-science/molecular-biology/cloning-and-expression/learning-center/guide-to-markers.html in Sigma-Aldrich.

Simmons, M.J., et al. 2012. NOTCH1 inhibition in vivo results in mammary tumor regression and reduced mammary tumorsphere-forming activity in vitro. Breast Cancer Research, 14(5): R126.

Smith, G.P. and Petrenko, V.A. 1997. Phage Display. Chemical Reviews, 97(2): 391-410.

Sommavilla, R., Lovato, V., Villa, A., Sgier, D. and Neri, D. 2010. Design and construction of a naïve mouse antibody phage display library. Journal of Immunological Methods, 353(1-2): 31-43.

Steinberger, P., Rader, C. and Barbas III, C. F. 2001. Analysis of Selected Antibodies. *In* Phage Display – A Laboratory Manual (C. F. Barbas III, D. R. Burton, J. K. Scott and G. J. Silverman eds), 1st ed., pp 11.1-11.24, Cold Spring Harbor Laboratory Press, New York

Stopeck, A., et al. 2012. The role of targeted therapy and biomarkers in breast cancer treatment. Clinical and Experimental Metastasis, 29(7): 807-19.

Suman, S., Das, T.P. and Damodaran, C. 2013. Silencing NOTCH signaling causes growth arrest in both breast cancer stem cells and breast cancer cells. British Journal of Cancer, 109(10): 2587-96.

Takebe, N., Nguyen, D. and Yang, S.X. 2014. Targeting Notch signaling pathway in cancer: Clinical development advances and challenges. Pharmacology and Therapeutics, 141(2): 140-149.

Thermo Fisher Scientific, Inc. 2008. 260/280 and 260/230 Ratios. Version 16 September 2014. http://www.nanodrop.com/Library/T009-NanoDrop%201000-and-NanoDrop%208000-Nucleic-Acid-Purity-Ratios.pdf in NanoDrop Products, http://www.nanodrop.com.

The Scripps Research Institute. 2014. Phage Display. Version 30 August 2014. http://www.scripps.edu/barbas/phage_display.html#pcomb3x in The Scripps Research Institute.

Tohidkia, M.R., Barar, J., Asadi, F. and Omidi, Y. 2012. Molecular considerations for development of phage antibody libraries. Journal of Drug Targeting, 20(3): 195-208.

UniProt. 2014. P80370- DLK1_HUMAN. Version 20-9-14. www.uniprot.org/uniprot/P80370 in UniProt.

van der Most, R. 2009. Combining Immunotherapy With Chemotherapy to Treat Cancer. V. 16-9-14. http://www.discoverymedicine.com/Robbert-G-Van-der-most/2009/07/25/combining-immunotherapy-with-chemotherapy-to-treat-cancer/ in Discovery Medicine, http://www.discoverymedicine.com.

Wang, J., Fu, L., Gu, F. and Ma, Y. 2011. Notch1 is involved in migration and invasion of human breast cancer cells. Oncology Reports, 26: 1295-303.

Watkins, N.A. and Ouwehand, W.H. 2000. Introduction to Antibody Engineering and Phage Display. Vox Sanguinis, 78(2): pp.72-9.

Welschof, M. et al. 2003. Generation and screening of a modular human scFv expression library from multiple donors. Methods in Molecular Biology, 207: 103-21.

WHO. 2014. Breast cancer: prevention and control. Version 27 July 2014. http://www.who.int/cancer/detection/breastcancer/en/index1.html in World Health Organization, http://www.who.int/en/.

Wu, Y., et al. 2010. Therapeutic ab targeting of individual Notch receptors. Nature, 464(7291): 1052-7.

Xin, L., et al. 2013. Human monoclonal antibodies in cancer therapy: a review of recent developments. Frontiers in Bioscience, 18: 765-72.

Yamabhai, M. 2014. Phage Display Technology. Version 13-9-14. http://www.sut.ac.th/iat/biotech/montarop/phd/introduction.htm in Suranaree University of Technology, http://web.sut.ac.th/2012/.

Yanai, H., et al. 2010. Dlk-1, a cell surface antigen on foetal hepatic stem/progenitor cells, is expressed in hepatocellular, colon, pancreas and breast carcinomas at a high frequency. The Journal of Biochemistry, 148(1): 85-92.

Zhu, Z. and Dimitrov, D.S. 2009. Construction of a Large Naïve Human Phage-Displayed Fab Library Through One-Step Cloning. Methods of Molecular Biology, 525: 129–42.

# 6. Appendix

## 6.1 Primers

**Table 6.1 - Primers used for HC Fd and VLCL fragments construction and Fab insert amplification**

| Region | Name | Sense | Sequence | Tails |
|---|---|---|---|---|
| VH | VH1 Fw | Sense | 5'CCAGCCTCGAGATGGCCCAGGTGCAGCTGG TGCAGTCTGG-3' | 3'-*XhoI* |
| | VH2 Fw | Sense | 5'CCAGCCTCGAGATGGCCCAGGTCAACTTGAA GGAGTCTGG3' | |
| | VH3 Fw | Sense | 5'CCAGCCTCGAGATGGCCGAGGTGCAGCTGG TGGAGTCTGG 3' | |
| | VH4 Fw | Sense | 5'CCAGCCTCGAGATGGCCCAGGTGCAGCTGC AGGAGTCGGG 3' | |
| | VH5 Fw | Sense | 5'CCAGCCTCGAGATGGCCCAGGTGCAGCTGTT GCAGTCTGG-3' | |
| | VH6 Fw | Sense | 5'CCAGCCTCGAGATGGCCCAGGTACAGCTGC AGCAGTCAGG-3' | |
| | JH 1/2 Rv | Anti-sense | 5'CGATGGGCCCTTGGTGGAGGCTGAGGAGAC GGTGACCAGGGTGCC3' | CH1 tail |
| | JH 3 Rv | Anti-sense | 5'CGATGGGCCCTTGGTGGAGGCTGAAGAGAC GGTGACCATTGTCCC 3' | CH1 tail |
| | JH 4/5 Rv | Anti-sense | 5'CGATGGGCCCTTGGTGGAGGCTGAGGAGAC GGTGACCAGGGTTCC3' | CH1 tail |
| | JH 6 Rv | Anti-sense | 5'CGATGGGCCCTTGGTGGAGGCTGAGGAGAC GGTGACCGTGGTCCC3' | CH1 tail |
| CH1 | IgG CH1 Rv | Anti-sense | 5'CGGCCCAGCCGGCCACTAGTGTCCACCTTG GTGCTGGGCTT3' | 5'-*SfiI*/*SpeI* |
| | CH1 Fw | Sense | 5' GCCTCCACCAAGGGCCCATCGGTC3' | |
| VLκ | VLκ1 Fw | Sense | 5'GGGCCCAGGCGGCCGAGCTCCAGATGACCC AGTCTCC-3' | 5'-*SfiI*/*SacI* |
| | VLκ1 Fw | Sense | 5'GGGCCCAGGCGGCCGAGCTCGTGATGACTC AGTCTCC3' | |
| | VLκ3a Fw | Sense | 5'GGGCCCAGGCGGCCGAGCTCGTGATGACAC AGTCTCC3' | |
| | VLκ3b Fw | Sense | 5'GGGCCCAGGCGGCCGAGCTCGTGATGACGC AGTCTCC3' | |
| | VLκ3c Fw | Sense | 5'GGGCCCAGGCGGCCGAGCTCGTGTTGACAC AGTCTCC3' | |
| | VLκ3d Fw | Sense | 5'GGGCCCAGGCGGCCGAGCTCGTGTTGACGC AGTCTCC3' | |
| | VLκ4 Fw | Sense | 5'GGGCCCAGGCGGCCGAGCTCGTGATGACCC ACACTCC-3' | |
| | VLκ5 Fw | Sense | 5'GGGCCCAGGCGGCCGAGCTCACACTCACGC GTCTCC3' | |
| | VLκ6 Fw | Sense | 5'GGGCCCAGGCGGCCGAGCTCGTGCTGACTC AGTCTCC-3' | |
| | VLκ Rv | Anti-sense | 5'GAAGACAGATGGTGCAGCCACAGT3' | CLκ tail |
| CLκ | CLκ Fw | Sense | 5'CGAACTGTGGCTGCACCATCTGTC 3' | |
| | CLκ Rv | Anti-sense | 5'GCGCCGTCTAGAATTAACACTCATTCCTGTTG AA-3' | 3'-*XbaI* |

| | | | | |
|---|---|---|---|---|
| **VLλ** | VLλ 1a Fw | Sense | 5'GGGCCCAGGCGGCCGAGCTCGTGTTGACGC AGCCGCCCTC 3' | 5'-*Sfi*I/*Sac*I |
| | VLλ 1b Fw | Sense | 5'GGGCCCAGGCGGCCGAGCTCGTGCTGACTC AGCCACCCTC 3' | |
| | VLλ 1c Fw | Sense | 5'GGGCCCAGGCGGCCGAGCTCGTGGTGACTC AGCCACCCTC 3' | |
| | VLλ 2 Fw | Sense | 5'GGGCCCAGGCGGCCGAGCTCGCCCTGACTC AGCCTCCCTCC3' | |
| | VLλ 3a Fw | Sense | 5'GGGCCCAGGCGGCCGAGCTCGAGCTGACTC AGCCACC-3' | |
| | VLλ 3b Fw | Sense | 5'GGGCCCAGGCGGCCGAGCTCGTGCTGACTC AGCCACC-3' | |
| | VLλ 4 Fw | Sense | 5'GGGCCCAGGCGGCCGAGCTCGTGCTGACTC AATCGCCCTC 3' | |
| | VLλ 6 Fw | Sense | 5'GGGCCCAGGCGGCCGAGCTCATGCTGACTC AGCCCCACTC 3' | |
| | VLλ 7/8a Fw | Sense | 5'GGGCCCAGGCGGCCGAGCTCGTGGTGACCC AGGAGCC3' | |
| | VLλ 7/8b Fw | Sense | 5'GGGCCCAGGCGGCCGAGCTCGTGGTGACTC AGGAGCC3' | |
| | VLλ 9 Fw | Sense | 5'GGGCCCAGGCGGCCGAGCTCGTGCTGACTC AGCCACCTTC3' | |
| | VLλ 10 Fw | Sense | 5'GGGCCCAGGCGGCCGAGCTCGGGCAGACTC AGCAGCT 3' | |
| | VLλ Rv | Anti-sense | 5'CGAGGGGGCAGCCTTGGGCTGACC3' | CLλ tail |
| **CLλ** | CLλ Fw | Sense | 5'GGTCAGCCCAAGGCTGCCCCC3' | |
| | CLλa Rv | Anti-sense | 5'GATTCTAGAGTAGGGGCCACTG3' | 3'-*Xba*I |
| | CLλb Rv | Anti-sense | 5'CATTCTAGAGCAGGGGCCACTG3' | |
| | pCOMB Fw | Sense | 5'AGTGGCACTGGCTGGTTTCGC3' | |
| | pCOMB Rv | Anti-sense | 5'CCATGGTGATGGTGATGGTGC3' | |

## 6.2 Libraries' CDR3 sequence diversity

**Table 6.2 – CDR3 sequence variability of κ Library tested Fab-containing clones.**

| κ Library | | | |
|---|---|---|---|
| **Clone** | **LC CDR3** | **Clone** | **HC CDR3** |
| K3 | LQDYNYPPT | K1 | AARIGAPATGFDS |
| K13 | MHALRTPPS | K10 | AHNSGHVVYYYYYMDV |
| K2 | MQATQFPPT | K26 | AHRRPRGWYVDFDY |
| K15 | MQGSHWPLT | K25 | AKVRSTYVRDTFDV |
| K26 | MQGSHWPYT | K6 | AQWRWGNWFDP |
| K14 | MQGTHWPAKDA | K13 | ARDLHEGHLSGWYLPFY |
| K1 | QLYGDSPMFT | K2 | ARDPGLRLGETSVFES |
| K8 | QLYGDSPMFT | K14 | AREGLGTTDRALAY |
| K18 | QLYGYSPLLM | K16 | AREITPDCSGGSCFFQSHGLDV |
| K9 | QQRSNWPPLT | K18 | ARGFESGGLYLGFDY |
| K21 | QQSFGTPWT | K8 | ARGYCSGGSCYSEAFDI |
| K6 | QQTFSFPWT | K15 | ARMKQLGGFYDFGLDV |
| K19 | QQYDTYPWT | K22 | ARQNVLTGYLYNYLDS |
| K17 | QQYFISPLT | K12 | ARRYDSGIDH |
| K22 | QQYGISPYT | K9 | ARTVGRATNFYCDN |
| K10 | QQYGNSPQT | K27 | ARVAYGNHFDF |
| K5 | QQYGSSPPFT | K4 | ARVPENFSYSSGSYRGPFDI |
| K25 | QQYGSSPRYT | K11 | ASSGRYRYDAFDF |
| K11 | QQYGSSPTT | K23 | TRRRNLATGNNGAPDFDY |
| K7 | QQYNNWPPYT | | |
| K24 | QQYNSYPWT | | |
| K23 | QQYVRSPYT | | |
| K16 | QQYYQTPYT | | |
| K20 | QQYYTIPLT | | |
| K4 | QVYGESPTWT | | |

**Table 6.3 - CDR3 sequence variability of λ Library tested Fab-containing clones.**

| λ Library | | | |
|---|---|---|---|
| Clone | LC CDR3 | Clone | HC CDR3 |
| L32 | AAWDDSLHGVI | L5 | AALAVAGTGYFDL |
| L14 | AAWDDSLKGVV | L24 | AHRRSGDGGEVERGAFDI |
| L17 | AAWDESLNGWV | L8 | AHSRGYCAANGCFSRGWSYNY |
| L26 | ALYMGLGIWV | L12 | AKVRMITMVRGDAFDI |
| L5 | ATWDVTLNGPNVV | L21 | ARAPVGDYSGFYTDSFDI |
| L11 | CSYAGSNSVI | L30 | ARATYGNHFDY |
| L6 | CSYAGSSTFEV | L13 | ARAVGLTYFDF |
| L28 | CSYALSRAI | L16 | ARDLPPYYGLSGNYYFDY |
| L3 | CSYVGGNSVVV | L19 | ARDRFQGGYSYGP |
| L29 | FSYTSRNTGL | L6 | ARDRVYDYGDYGSMDV |
| L1 | GTWDSSLGVVL | L9 | ARDRVYDYGDYGSMDV |
| L4 | LLSYGGAWV | L18 | ARERRQLGVYYYGMDV |
| L8 | LLSYSGGRYVDVI | L11 | ARGRLQLTRRPTAIRYTPFDV |
| L22 | QSYDNSLSGSWV | L26 | ARMKQLGGFYDFGLDV |
| L25 | QSYDSSNRSWV | L15 | ARRTVTTPGFDC |
| L2 | QSYDSSNVV | L10 | ARVASSGYFSDY |
| L20 | QTWATDGVS | L7 | ARVPQNDYGDYYYFEN |
| L18 | QTWGTGGVA | L28 | ARWLQCFDY |
| L7 | QVWDSSSDHVV | L17 | ASSGRYRYDAFDF |
| L31 | SAWDDRLSGHVV | L23 | ASSGRYRYDAFDF |
| L24 | SSYAGPAEV | L27 | ATSRGGRNAFDI |
| L15 | SSYAGSSTSEV | L3 | VHPKSKGPFYDKPVNAFHV |
| L21 | SSYTDSSTFIVV | L4 | VKESEFGYYRTADY |
| L30 | VLYMGSGIPV | L14 | VRIYYQYGDYPHFHYMDV |
| | | L1 | VRNKYYYGSGSRRGRVDGFDV |

## 6.3 D2/D19 clone sequence

### A) Nucleotide sequence

#### A.1) LC insert

- Length: 634 bp
- VBASE2 sequence analysis:

Legend: - FR1; CDR1; FR2; CDR2; FR3; CDR3; FR4

GTGCTGACTCATCGCCCTCTGCCTCTGCCTCCCTGGGAGCCTCGGTCAATTTCACCTGCACTCTGGGC
AGTGGACACAAAAACTTCGCCATCGCATGGCATCAGCAGCAGCCAGACAAGGGCCCTCGCTACCTGAT
GAACCTTGATAGTGAAGGCAGCCATACGAAGGGGGACGGGATCCCTGACCGCTTCTCAGGCTCCAGCG
CAGGGGCTGATCGCTACCTCACCATCTCCAGCCTCCAGTCTGAAGATGAGGCTGTCTATTACTGTCAG
ACCTGGGCCACCGACGGTGTGTCATTCGGCGGAGGGACCAAGCTGACCGTCCTAAGTCAGCCCAAGGC
TGCCCCCTCGGTCACTCTGTTCCCACCCTCCTCTGAGGAGCTTCAAGCCAACAAGGCCACACTGGTGT
GTCTCATAAGTGACTTCTACCCGGGAGCCGTGACAGTGGCCTGGAAGGCAGATAGCAGCCCCGTCAAG
GCGGGAGTGGAGACCACCACACCCTCCAAACAAAGCAACAACAAGTACGCGGCCAGCAGCTACCTGAG
CCTGACGCCTGAGCAGTGGAAGTCCCACAGAAGCTACAGCTGCCAGGTTACGCATGAAGGGAGCACCG
TGGAGAAGACAGTGGCCCCTGC

#### A.2) HC insert

- Length: 797 bp
- VBASE2 sequence analysis:

Legend: - FR1; CDR1; FR2; CDR2; FR3; CDR3; FR4

CCCTTGTCCAGGGGCCTGTCGCACCCAGTGCATATAGTAGTCGGTGATGCTGTATCCAGAAGCCTTGC
ATGAGACCTTCACTGAGGCCCCAGGCATCTTCACCTCAGCCCCAGACTGCAACAGCTGCACCTGGGCC
ATCTCGAGATGGCCGAGGTGCAGCTGGTGGAGTCTGGGGGAGGCGTGGTCCAGCCTGGGAGGTCCCTG
AGACTCTCCTGTGCAGCGTCTGGATTCACCTTCAGTAGCTATGGCATGCACTGGGTCCGCCAGGCTCC
AGGCAAGGGGCTGGAGTGGGTGGCAGTTATATGGTATGATGGAAGTAATAAATACTATGCAGACTCCG
TGAAGGGCCGATTCACCATCTCCAGAGACAATTCCAAGAATACGCTGTATCTACAAATGAACAACCTG
AGAGGCGAGGACACGGCTGTCTACTACTGTGCGAGGGATGCCAACAGCTGGTACAGTGGCAATTACTT
TGACCACTGGGGCCAGGGAACCCTGGTCACCGTCTCTTCAGCCTCCACCAAGGGCCCATCGGTCTTCC
CCCTGGCACCCTCCTCCAAGAGCACCTCTGGGGGCACAGCGGCCCTGGGCTGCCTGGTCAAGGACTAC
TTCCCCGAACCGGTGACGGTGTCGTGGAACTCAGGCGCCCTGACCAGCGGCGTGCACACCTTCCCGGC
TGTCCTACAGTCCTCAGGACTCTACTCCCTCAGCAGCGTGGTGACCGTGCCCTCCAGCAGCTTGGGCA
CCCAGACCTACATCTGCAACGTGAATCACAAGCCCAGCACCAAGGTGAC

## B) Amino Acid Sequence

#### B.1) VL region

xARADSSPSASASLGASVNFTCTLGSGHKNFAIAWHQQQPDKGPRYLMNLDSEGSHTKGDGIPDRFSG
SSAGADRYLTISSLQSEDEAVYYCQTWATDGVSFGGGTKLTVL

#### B.2) VH region

EVQLVESGGGVVQPGRSLRLSCAASGFTFSSYGMHWVRQAPGKGLEWVAVIWYDGSNKYYADSVKGRF
TISRDNSKNTLYLQMNNLRGEDTAVYYCARDANSWYSGNYFDHWGQGTLVTVSS