

**IDENTIFICATION AND UTILIZATION OF
AUTOLOGOUS ANTI-TUMOR ANTIBODIES
FOR THE DIAGNOSIS AND PROGNOSIS OF CANCER**

**A THESIS SUBMITTED TO
THE GRADUATE SCHOOL OF ENGINEERING AND SCIENCE
OF BILKENT UNIVERSITY
IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR
THE DEGREE OF
DOCTOR OF PHILOSOPHY
IN
MOLECULAR BIOLOGY AND GENETICS**

By
Şükrü Atakan
December 2015

IDENTIFICATION AND UTILIZATION OF AUTOLOGOUS ANTI-TUMOR
ANTIBODIES FOR THE DIAGNOSIS AND PROGNOSIS OF CANCER
By Şükrü Atakan
December 2015

We certify that we have read this thesis and that in our opinion it is fully adequate,
in scope and in quality, as a thesis for the degree of Doctor of Philosophy.

Ali Osmay Güre (Advisor)

Özlen Konu

Serkan Göktuna

Uğur Özbek

Eda Çelik Akdur

Approved for the Graduate School of Engineering and Science:

Levent Onural
Director of the Graduate School

ABSTRACT

IDENTIFICATION AND UTILIZATION OF AUTOLOGOUS ANTI-TUMOR ANTIBODIES FOR THE DIAGNOSIS AND PROGNOSIS OF CANCER

Şükrü Atakan

Ph.D. in Molecular Biology and Genetics

Advisor: Ali Osmay Güre

December 2015

Lung cancer is the leading cause of cancer related death worldwide. Current diagnostic methods have limited power and unable to extend patient life significantly. SCLC; the most aggressive subtype of lung cancer is an immunogenic cancer type and able to elicit an immune response of which autologous antibodies are a measurable component. These antibodies are elicited even when the tumor is microscobic and impossible to be diagnosed clinically by the current methods of diagnosis thus antibodies can be utilized for early diagnosis. We aimed to develop a method to identify novel autologous antibodies, identify these antibodies for SCLC, Colorectal, Gastric and Ovarian cancers and validate these antibodies for SCLC diagnosis and prognosis and investigate their utility for autoimmune disease.

We have developed and optimized PA screening for novel autologous antibody discovery. We have screened PA with serum pools of cancer patients (SCLC, Colorectal, Gastric and Ovarian), BD and healthy controls since PAs have many advantages compared to other discovery methods like SEREX. We have also performed sensitivity and specificity evaluations by screening custom PAs by individual sera. Image analysis softwares developed by collaboration utilized for evaluation of the screenings. The filtered valuable clones were ordered from the PA manufacturer and HisTagged protein expression and purification was performed with these clones. Pure proteins were screened with 3 independent SCLC and 2 Healthy control cohorts by an iterative ELISA approach for validation of these antibodies as valuable biomarkers. ELISA results were also confirmed by Western blotting. Monte Carlo, SVM and PC were utilized for cut-off determination, panel formation and ROC plotting. AUC was compared for evaluation of diagnostic power. Kaplan-Meier, UCR and MCR analysis was performed for prognostic analysis of the valuable antibodies. Separately protein expression and autologous antibody presence correlation was evaluated by comparison of IHC and ELISA. The same autologous antibody identification strategy was utilized as a collaborative support for an independent study for identification of NBD specific biomarkers.

We have identified 23 distinct autologous antibody biomarkers for SCLC after evaluation of PA and custom PA screenings. For 8 of these antibodies we have completed ELISA screening for all 3 SCLC and 2 healthy control cohorts. 6 of these autologous antibodies were shown to be valuable as a panel for SCLC diagnosis both by MC and SVM. Utilization of 4 of these antibodies; SOX2, p53, POLB and C11orf20, as a panel resulted in superior AUC thus high sensitivity and specificity values (55% sensitivity, 90% specificity). PC method resulted in higher AUC even only by combination of SOX2 and p53 (82% sensitivity, 90% specificity). Although individual correlations were identified, we were unable to show a significant correlation of seropositivity with survival for any of the antibodies which is common to all cohorts. We have identified a significant correlation between SOX2 antigen expression intensity and autologous antibody presence. Mtch1 was identified as a NBD specific autologous antibody by the utilization of our autologous antibody discovery and validation methodology.

We were able to identify a panel of 4 antibodies; SOX2, p53, POLB and C11orf20, which resulted in 55% sensitivity at 90% specificity for SCLC. 2 of these antibodies were identified by this study as novel biomarkers; POLB and C11orf20. The panel is capable of exceeding the diagnostic power of the only commercially available diagnostic kit; EarlyCDT-Lung. PC method is very promising since a sensitivity value of 82% was reached at 90% specificity which is a diagnostic power comparable that of low-dose CT. As a future perspective we are planning to apply PC method to all the PA data and develop a kit based on this method to be utilized for SCLC diagnosis.

Keywords: SCLC, protein array, diagnosis, prognosis, autologous antibodies, biomarker, ELISA, sensitivity, specificity.

ÖZET

KANSERİN TEŞHİS VE PROGNOZUNDA KULLANILMAK ÜZERE OTOLOG ANTI-TÜMÖR ANTİKORLARIN BELİRLENMESİ VE KULLANIMI

Şükrü Atakan

Moleküler Biyoloji ve Genetik, Doktora

Tez Danışmanı: Ali Osmay Güre

Aralık 2015

Akciğer kanseri dünya çapında en çok ölüme yol açan kanserdir. Mevcut teşhis yöntemleri kısıtlı güce sahiptir ve hasta ömrünü anlamlı olarak uzatamamaktadır. Küçük hücreli akciğer kanseri (KHAK); en agresif akciğer kanseri alt türü olarak bağışıklık yanıtına yol açmaktadır ve bu yanıtın ölçülebilir ajanlarından biri otolog anti-tümör antikordur. Bu otolog antikordur tümör henüz mikroskobik aşamadayken dahi ortaya çıkmaktadır ve bu sayede diğer teşhis yöntemlerinin tümörü tespit edemeyeceği durumlarda dahi tümör varlığını erken aşamada teşhis etmeye yarayabilir. Çalışmamızda otolog anti-tümör antikordur keşfetmek için yeni bir yöntem geliştirmeyi, KHAK, Kolorektal, Mide ve Over kanserlere özgü otolog antikordur keşfetmeyi, KHAK teşhisi ve prognozunda yararlı olabilecek otolog anti-tümör antikordur belirlenmeyi ve doğrulamayı ve otolog antikordur otoimmün hastalık için yararlılığını araştırmayı hedefledik.

Çalışmamızda bilinmeyen otolog anti-tümör antikordur keşfetmek için PA taraması temelli bir yöntem geliştirdik ve optimize ettik. Kanser (KHAK, kolorektal, mide ve over) hastaları, Behçet hastalığı ve sağlıklı kontrollerden oluşturulan serum havuzları ile PA taraması gerçekleştirdik. PA taramasının SEREX gibi diğer otolog antikör teşhis yöntemlerine göre üstün yönleri vardır. Havuz taramalarından seçilmiş klonlardan oluşan arraylerin tek tek serumlar ile tekrar taranması ile hassasiyet ve özgüllük değerlendirmeleri gerçekleştirdik. Kolaborasi ile oluşturulmuş görüntü işleme yazılımları tarama çalışmalarının sonuçlarının değerlendirilmesinde kullanılmıştır. Değerlendirmeler sonucu seçilen klonlar PA üreticisinden temin edilmiştir ve bu klonlar ile His ile işaretlenmiş protein üretimi ve saflaştırması gerçekleştirilmiştir. Saf proteinler ile 3 farklı KHAK ve 2 farklı kontrol grubu için sonuçların her defasında tekrar değerlendirildiği ELISA taramaları yapılmıştır. Bu taramalardaki amaç belirlenen otolog antikordur değerli birer biyobelirteç olduğundan emin olmaktır. ELISA sonuçlarımız Western blot yöntemiyle doğrulanmıştır. MC, SVM ve PC yöntemleri eşik değeri belirlemek, panel oluşturmak ve ROC eğrisi çizimi için uygulanmıştır. Teşhis gücünün kıyaslanması için AUC değerleri analiz edilmiştir. Kaplan-Meier, UCR ve MCR yöntemleri antikordur varlığı ile yaşam süresi arasındaki ilişkinin değerlendirilmesi için uygulanmıştır. Ayrıca protein ifadesi ve otolog antikör varlığı arasındaki ilişki IHC ve ELISA değerlerinin

analizi ile araştırılmıştır. Otolog antikör tespiti için geliştirdiğimiz yöntem ayrıca kolaboratif bir destek olarak NBD'ye özgü biyobelirteçlerin bulunmasında uygulanmıştır.

KHAK için 23 farklı otolog antikör biyobelirteç PA taramaları sonucunda anlamlı olarak seçilmiştir. Bu antikörlerin 8'i için 3 KHAK ve 2 sağlıklı grubu için tüm ELISA taramaları tamamlanmıştır. Bunların altısından oluşturulacak bir panelin değerli olacağı MC ve SVM ile gösterilmiştir. Bu 6 antikörün dördünden oluşturulan panelin; SOX2, p53, POLB ve C11orf20, AUC değeri tüm 6 antikör ile oluşturulan AUC değerinden çok farklı değildir ve yüksek hassasiyet ve özgüllük elde edilebilmektedir (55% hassasiyet, 90% özgüllük). PC yöntemi ile yalnızca SOX2 ve p53 içeren panel ile oluşturulan AUC değeri diğer tüm AUC değerlerinden daha yüksektir (82% hassasiyet, 90% özgüllük). Tek başına değerlendirildiğinde genel yaşam süresi ile istatistiksel olarak anlamlı şekilde ilişkili bulunmasına rağmen hiçbir antikör varlığı tüm gruplar için genel yaşam süresi ile ilişkili bulunmamıştır. Çalışmamızda SOX2 protein ifade yoğunluğu ile otolog antikör varlığı arasında istatistiksel anlamlı ilişki bulduk. Mtch1'in NBD'ye özgü bir otolog antikör olarak keşfi ve doğrulanması ile yöntemimizin otoimmün hastalıklarda da yararlılığı anlaşılmıştır.

Belirlediğimiz 4 otolog antikör ile oluşturduğumuz panel; SOX2, p53, POLB ve C11orf20, KHAK teşhisi için 55% hassasiyette 90% özgüllüğe sahiptir. Bu otolog antikörlerin ikisinin ilk defa bu çalışma ile SCLC teşhisinde kullanılabilirliği gösterilmiştir; POLB ve C11orf20. Dört antikör panel akciğer kanseri için piyasada mevcut tek otolog antikör temelli kit olan EarlyCDT-Lung'dan daha yüksek teşhis gücüne sahiptir. PC yöntemi ile 90% özgüllük için 82% hassasiyet değerine ulaşılabilir. Bu teşhis gücü düşük doz bilgisayarlı tomografinin teşhis gücüne eşdeğerdir. Gelecek planı olarak PC yöntemini şimdiye kadar oluşturduğumuz tüm PA sonuçlarına uygulamayı ve bu yöntem temelli çalışan bir KHAK teşhis kiti geliştirmeyi hedefliyoruz.

Anahtar sözcükler: KHAK, protein array, teşhis, prognoz, otolog antikörler, biyobelirteç, ELISA, hassasiyet, özgüllük.

ACKNOWLEDGEMENTS

I would like to express my gratitude to my advisor Assist. Prof. Dr. Ali Osmay Güre for his invaluable contributions to this work and to my personal development. I appreciate so much his sincere and honest attitude, his patience, always encouraging positive criticisms and his influential didactic style. He gave me the chance to learn by trial and error which is very valuable for me. He has also been very supportive for my extracurricular activities in context of entrepreneurship thus gave me an opportunity to discover novel possibilities for my career besides autologous antibodies.

I would like to thank Assist. Prof. Dr. Özlen Konu and Prof. Dr. Uğur Özbek for their constructive and positive criticisms about my Ph.D. project.

I would like to thank Alper Poyraz for his invaluable contributions to the project and his support, encouragements and executive personality. I also appreciate his partnership in business by his positive, energetic and influential mood.

I would like to thank Serhat Yıldırım for his tremendous efforts for the project. He helped the project by providing hard work and performing attention and precision requiring, labor intensive and long lasting experiments.

I would like to thank Prof. Dr. Burçak Vural for her contributions to this project, her encouragements, positive attitude and her friendship. I would like to thank Elif Uğurel, Filiz Çavuş and Elçin Şehitoğlu for their friendships and contributions to the project.

I would like to thank Prof. Dr. Burçak Vural, Dr. Funda Demirağ, Dr. Hülya Bayız and Prof. Dr. Pınar Saip for providing serum samples and the clinical data.

I would like to thank Assoc. Prof. Dr. Çiğdem Gündüz Demir and Can Fahrettin Koyuncu for their collaboration in image processing software development.

I would like to thank former and present members of AOG Lab; Alper, Barış, Kerem, Sinem, Seçil, Waqas, Mehdi, Zeynep, Murat, Duygu, Aydan, Esen, Rasim and Derya for their kind contributions to the project and their friendship.

I would like to thank the senior students who have contributed to this project; Alper, Serhat, Deniz, İpek, Berrak, Dilce, Fatma Zehra, Onur, Deniz, Duygu, Anıl, Çağrı, Ferhat, Erdi, Ezgi and Gökhan.

I would like to thank Alper and Barış specially for their efforts for contributing my extracirricular R&D activities during their undergraduate study.

I would like to thank all my instructors in Bilkent MBG for their contributions to my scientific development. It is a honor for me to be a graduate of Bilkent MBG and doing my PhD as well and Bilkent's contributions to my career as well as my philosophy about life.

I would like to thank to my friends; Burak, Duygu, Ender, Ceyhan, Kerem, Sinem, Gurbet, Nilüfer, Zeynep, Gizem and Aydan for our wonderful memories.

I would like to thank all the members of the Bilkent MBG family as a community for their contributions to my knowledge, my personal development and the warm family environment.

I would like to thank Rengül Hoca and Volkan Hoca for accepting me as a member of their family during my PhD study and giving me the opportunity to meet my wife.

I would like to thank my wife, Işıl for her love, her mental and emotional support and her patience during my PhD.

I would like to thank my family for all their support and encouragement about my education and my decisions.

I would like to thank to TÜBİTAK for the BİDEB 2241 scholarship during my PhD.

I would like to thank to PRZ BioTech and Sentegen for employing me and encouraging and supporting my PhD studies in parallel to their employment.

I dedicate my thesis firstly to my advisor Assoc. Prof. Dr. Ali Osmay Güre and all my instructors whose lectures I have attended during my undergraduate and PhD studies at Bilkent University. They have tremendously contributed to my scientific and philosophical development.

Contents

1	INTRODUCTION.....	1
1.1	LUNG CANCER AND IT'S DIAGNOSIS.....	1
1.1.1	Lung Cancer Risk factors, Incidence and Mortality	1
1.1.2	Current methods for dignosis of lung cancer	1
1.1.3	Potential of biomarkers for lung cancer diagnosis / early diagnosis.....	2
1.2	IMMUNE SYSTEM AND CANCER	2
1.2.1	Immune response against cancer	2
1.2.2	Serum reactivity against cell surface antigens	3
1.2.3	Serum reactivity against intracellular antigens	3
1.2.4	Methylcholanthrene induction causes tumor cells to be rejected by the host.....	3
1.2.5	Immunoediting is a driving force for less imunogenic tumor development	4
1.2.6	Autologous CTLs lyse antigen presenting tumor cells.....	4
1.2.7	Autologous antibodies as part of the immune response to cancer	4
1.2.8	Autologous antibody elucidation mechanisms	5
1.2.9	Antigen expression and autologous antibody elucidation	6
1.2.10	Autologous antibodies and PNS.....	6
1.2.11	Autologous antibodies in SCLC with/without PND.....	7
1.3	DIAGNOSIS AND EARLY DIAGNOSIS OF CANCER	8
1.3.1	Autologous antibodies precedes PND and clinical cancer diagnosis.....	8
1.3.2	p53 antibody level stability / fluctuations can predict cancer development	9
1.3.3	Sensitivity and specificity of autologous antibodies	10
1.3.4	Autologous antibodies already in clinical use for diagnosis of cancer	10
1.4	AUTOLOGOUS ANTIBODY DISCOVERY METHODS	10
1.4.1	Methods for autologous antibody discovery	10
1.4.2	Limitations of SEREX	13
1.4.3	Theoretical comparison of SEREX with PAs	15
1.5	AUTOLOGOUS ANTIBODIES AND PROGNOSIS.....	16
1.5.1	Autologous antibodies can be utilized for prognosis	16
1.5.2	Autologous antibody positive sera have cytotoxic effect on cancer cell line.....	16
1.5.3	Autologous antibody positivity correlates with <i>in-vivo</i> cytotoxicity	16
1.5.4	Autologous antibodies and survival.....	17
1.5.5	Autologous anti-tumor antibodies identified by our previous studies	17

1.5.6	Anti-SOX1 autologous antibodies are specific to cancer not PND	17
1.5.7	SOX2 antibodies and prognosis of multiple myeloma (MM)	17
1.5.8	SOX2 and ZIC2 antibodies and survival in SCLC.....	18
1.6	SENSITIVITY AND SPECIFICITY OF PANELS	19
1.6.1	Diagnosis by evaluating a panel of antibodies	19
1.6.2	Sensitivity and Specificity of already identified autologous antibodies	20
1.6.3	Predicted Sensitivity and Specificity of novel autologous antibodies	22
1.6.4	Commercially available diagnostic technologies	23
2	AIM, RELEVANCE AND RATIONALE	24
2.1	AIMS.....	24
2.2	RELEVANCE	25
2.3	RATIONALE.....	25
3	STUDY DESIGN	27
4	MATERIALS AND METHODS	28
4.1	METHOD DEVELOPMENT.....	28
4.1.1	Optimization Experiments for PA screening protocol determination	28
4.1.2	Serum Selection for Pooling.....	28
4.1.3	Serum Dilution	28
4.1.4	Pool Dilution	29
4.1.5	2° Antibody dilution.....	29
4.1.6	Q-Dot dilution	29
4.1.7	Strategy for autologous antibody discovery	29
4.1.8	Protein Arrays	31
4.2	NOVEL AUTOLOGOUS ANTIBODY IDENTIFICATION	32
4.2.1	Screening Strategy and Cohorts.....	32
4.2.2	Patient and Healthy Control Cohorts	32
4.2.3	PA Screening	33
4.2.4	Clone Selection Strategy based on PA Screening Results	34
4.2.5	Manual Selection Software (MSS) Development for selection of signals from PA screening results	34
4.2.6	Signal Analysis Software (SAS) and Pixel Count (PC) Method for Numerical Analysis.....	35
4.2.7	Custom PA Screening.....	37
4.2.8	Custom PA screening results' evaluations	38

4.3	INTER-ASSAY VALIDATION	38
4.3.1	Insert analysis of pQE plasmids	38
4.3.2	Clone protein expression analysis by IPTG induction and SDS-PAGE	38
4.3.3	HisTagged Protein Expression and Purification.....	39
4.3.4	ELISA	39
4.3.5	ELISA Data Normalization	40
4.3.6	Sensitivity and Specificity calculation	40
4.3.7	ELISA Evaluations	41
4.3.8	Western blotting.....	41
4.3.9	Optimum cut-off determination for ELISA results by MC Method.....	42
4.3.10	Optimum antigen panel determination from Custom Array Screening Evaluations by MC Method	42
4.4	PROGNOSTIC ANALYSIS OF AUTOLOGOUS ANTIBODIES.....	43
4.4.1	R packages utilized for diagnostic and prognostic analysis.....	43
4.5	ANTIGEN EXPRESSION AND ANTIBODY CORRELATION	43
4.5.1	Patient and control population.....	43
4.5.2	Enzyme-linked immunosorbent assay	44
4.5.3	Western analysis of SOX2 seroreactivity	44
4.5.4	Immunohistochemistry.....	44
4.5.5	Statistical analysis.....	45
4.6	AUTOLOGOUS ANTIBODIES IN AUTOIMMUNE DISEASE	45
4.6.1	Identification of autologous antibodies in NBD	46
4.6.2	ELISA and Western blot validation of autologous antibody levels.....	46
5	RESULTS	47
5.1	METHOD DEVELOPMENT.....	47
5.1.1	Serum selection for pooling.....	47
5.1.2	Serum dilution	48
5.1.3	Pool dilution	49
5.1.4	2° Ab dilution.....	50
5.1.5	Q-Dot dilution.....	51
5.2	NOVEL AUTOLOGOUS ANTIBODY IDENTIFICATION	51
5.2.1	Quality Control of PA Screening.....	51
5.2.2	Clone selection from pool screenings	53
5.2.3	Clones from Testis pool screening	55

5.2.4	Clones from FB pool screening	56
5.2.5	Sensitivity and Specificity of custom PA results	57
5.3	INTER-ASSAY VALIDATION	62
5.3.1	Insert Analysis of pQE plasmids	62
5.3.2	IPTG induction control of the selected clones	64
5.3.3	Iterative ELISA Approach	65
5.3.4	Stepwise clone selection by ELISA evaluations	66
5.3.5	Normalization for ELISA data	67
5.3.6	Coefficient of Variation (CV) Analysis	68
5.3.7	Quality control for serum integrity	69
5.3.8	ELISA Data.....	70
5.3.9	Manual cut-off determination	72
5.3.10	Cut-off determination by Monte Carlo	73
5.3.11	ELISA and Western blot comparison.....	74
5.3.12	ELISA and custom PA comparison.....	76
5.3.13	Custom PA and Western blot comparison	78
5.3.14	Sensitivity and Specificity analysis	79
5.3.15	ROC analysis	81
5.3.16	Accuracy of antibody combination	83
5.3.17	Comparison of evaluation methods.....	84
5.4	PROGNOSTIC ANALYSIS OF AUTOLOGOUS ANTIBODIES.....	88
5.5	ANTIGEN EXPRESSION AND ANTIBODY CORRELATION	100
5.6	AUTOLOGOUS ANTIBODIES IN AUTOIMMUNE DISEASE	104
5.6.1	Autologous antibody identification for NBD	104
6	SUMMARY.....	107
6.1	Significance of the study.....	107
6.2	Major findings.....	108
7	DISCUSSION AND CONCLUSION	112
7.1	Method Development and Protein Arrays	112
7.2	PA Screening Evaluations.....	113
7.3	Contradictions between PA and custom PA Screening Evaluations	114
7.3.1	Clones not recognized by pool but recognized by individual sera	114
7.3.2	Clones recognized by pool but not recognized by individual sera	114

7.4	Novel autologous antibody identification and their diagnostic value	114
7.5	The literature survey for the most valuable 6 antigens	115
7.5.1	SOX2	116
7.5.2	p53.....	116
7.5.3	C11orf20	116
7.5.4	POLB	117
7.5.5	ALMS1.....	118
7.5.6	TUSC1	119
7.6	Prognostic analysis of autologous antibodies	119
7.7	Correlates of antigen expression, presence of autologous antibodies and clinical parameters	120
7.8	Evaluation of autologous antibodies in autoimmune disease	122
8	FUTURE PERSPECTIVES.....	124
8.1	Application of PC to the PA data	124
8.2	Diagnostic kit development.....	124
8.3	Colorectal, Gastric and Ovarian cancers.....	125
8.4	Utilization of antibodies for early diagnosis of cancer	125
9	BIBLIOGRAPHY	126

List of Figures

Figure 1. Conventional techniques for novel autologous antibody discovery and validation	11
Figure 2. Anti-SOX2, HuD, CRMP5, p53, NY-ESO-1 and/or recoverin autologous antibodies in SCLC diagnosis.....	19
Figure 3. Anti-SOX2, HuD, CRMP5, p53, NY-ESO-1 and/or recoverin autologous antibodies in NSCLC diagnosis.....	20
Figure 4. Complete experimental strategy for discovery and validation of novel autologous antibodies	30
Figure 5. PA format and a representative positive signal.....	31
Figure 6. PA Spotting pattern is composed of 12 proteins in duplicate and a reference InkDot. ..	32
Figure 7. Representative output of MSS	35
Figure 8. Evaluation method of PC	37
Figure 9. Non-specific background staining by some SCLC sera (left) and their corresponding ELISA values (right)	47
Figure 10. SCLC sera causing undesirable background staining.....	48
Figure 11. Final serum dilution ratio determination for ideal signal / background ratio	49
Figure 12. Final serum dilution ratio determination by pool dilution	50
Figure 13. 2 ^o Ab dilution ratio comparison.....	51
Figure 14. Q-Dot dilution ratio comparison.....	51
Figure 15. Repeats of custom PA screening by the same serum for reproducibility assessment ..	52
Figure 16. Valuable clones selected from Testis pool screenings and the overlap for different cancers.....	53
Figure 17. Valuable clones selected from FB pool screenings and the overlap for different cancers.....	54
Figure 18. Our iterative ELISA strategy for identification of best antigens by performing the minimum number of screening	65
Figure 19. Stepwise antigen selection strategy by utilizing standart deviation cut-offs	66
Figure 20. ELISA data normalization for plate and serum	67
Figure 21. Quality control for the serum antibody integrity	69
Figure 22. Scatter dot plot of ELISA values of the most valuable 6 antigens	71
Figure 23. Manual cut-off determination for seropositivity for the ELISA results	72
Figure 24. Cut-off determination utilizing the MC algorithm.....	73

Figure 25. Confirmation of SOX2 ELISA with Western blot for a representative group of sera.....	74
Figure 26. ELISA and Western blot correlation for a group of patients for the 6 antigens	75
Figure 27. ELISA and custom PA correlation	77
Figure 28. Correlation of PA with Western blotting for representative clones.....	78
Figure 29. SVM and MC comparison for ELISA values	82
Figure 30. Feature selection for the antigen panel accuracy evaluation	83
Figure 31. Comparison of SAS and ELISA for ROC by MC	85
Figure 32. ELISA, PC by MC and Chest X-ray and Low-dose CT comparison for ROC	87
Figure 33. Seropositivity and survival correlation for different cut-offs	94
Figure 34. Volcano plots; seropositivity and survival correlation for SCLC cohorts.....	97
Figure 35. Seropositivity and overall survival correlation by Kaplan-Meier graphs.....	98
Figure 36. SOX2 autologous antibody and SOX2 expression correlation.....	101
Figure 37. SOX2 protein expression and autologous antibody correlation analysis	103
Figure 38. Scatter dot plots of Mtch1 antibodies in the cohorts.....	104

List of Tables

Table 1: Autologous antibodies in cancer and PND	7
Table 2: p53 antibody levels in case of cancer development and controls	9
Table 3: Correlation between SOX2 phage and ELISA results	12
Table 4: Technical limitations of SEREX (A)	13
Table 5: Technical limitations of SEREX (B).....	14
Table 6: mRNA representation limitations of SEREX	14
Table 7: Summary of limitations of SEREX.....	15
Table 8: Sensitivity and specificity of anti-tumor antibodies in lung cancer.....	21
Table 9: Sensitivity values of autologous antibodies useful in SCLC diagnosis at 100% specificity	21
Table 10: Estimated effect of novel autologous antibodies in sensitivity of SCLC diagnosis	22
Table 11: Sensitivity and specificity values of autologous anti-tumor antibodies for other solid tumors.....	22
Table 12: Representative output of SAS	36
Table 13: List of clones ordered from Testis pool screening.....	55
Table 14: List of clones ordered from FB pool screening.....	56
Table 15: Final list of clones screened with ELISA.....	61
Table 16: CV for ELISA data as quality control (A: Medium Signals).....	68
Table 17: CV for ELISA data as quality control (B: High Signals)	68
Table 18: Individual sensitivity values of the most valuable 6 antigens in 100%-90% specificity range by MC	79
Table 19: Sensitivity and specificity values of the antigen panels for the best MC cut-offs	80
Table 20: Univariate Cox regression analysis for overall survival	88
Table 21: Antibody frequency and overall survival (cut-off: Mean + 3 * StDev).....	90
Table 22: Cox Regression analysis and overall survival.....	92
Table 23: Final list of valuable clones after custom PA screening of BD	105
Table 24: Most valuable MC panels for BD and their antibody content	106

Abbreviations

AUC	Area under the curve
BD	Behçet's Disease
CT	Computerized Tomography
CV	Coefficient of Variation
ELISA	Enzyme Linked Immunosorbent Assay
FB	Fetal Brain
IHC	Immunohistochemistry
MC	Monte Carlo
MCR	Multivariate Cox Regression
MSS	Manual selection software
NBD	Neuro Behçet's Disease
NSCLC	Non-small cell lung cancer
PA	Protein Array
PC	Pixel Count
PNPP	p-Nitrophenyl Phosphate
SAS	Signal analysis software
SCLC	Small cell lung cancer
SD	Standard Deviation
SOX2	SRY (Sex Determining Region Y)-Box2
SVM	Support Vector Machine
UCR	Univariate Cox Regression

Chapter 1

1 INTRODUCTION

1.1 LUNG CANCER AND IT'S DIAGNOSIS

1.1.1 Lung Cancer Risk factors, Incidence and Mortality

Lung Cancer caused 1.1 million male and 491.200 female deaths in 2012 as the primary cause of cancer related death worldwide [1]. Tobacco smoking is the main reason of Lung cancer, responsible for 90% of the lung cancers [2, 3], and the other causes are radon exposure, asbestos exposure, passive smoking, arsenic and polycyclic hydrocarbons [3]. The main reason why Lung cancer is so deadly is that it is detected mostly when the disease is disseminated. SCLC constitutes 20% of all lung cancer cases [4, 5] and is the most aggressive subtype. Overall 5-year survival rate for SCLC is only 6% [3].

One of the highest number of yearly new cancer cases belongs to lung cancer among all cancers and the highest number of cancer related mortality belongs to lung cancer. The earlier stages of lung cancers correlate with good overall survival compared to the later stages. The biggest problem about lung cancer is that the disease doesn't cause any symptoms at early stages and at later stages any type of therapy doesn't cure the cancer. The survival rate of lung cancer drops from 37% for the first year to 8% for the 5. year [6]. So that early novel diagnosis methods for Lung Cancer are urgently required.

1.1.2 Current methods for dignosis of lung cancer

Currently the method used for lung cancer diagnosis is CT [4]. More than 60% of patients are diagnosed at advanced stages when a cure is not possible [7]. SCLC patients have 5-year relative survival rates of 31% for Stage 1, 19% for Stage 2, 8% for Stage 3 and only 2% for Stage 4 [1]. Since early diagnosis can increase the overall survival of SCLC patients [8-10] a novel method which can be practical, non-invasive and cheap is urgently required.

Currently, after presentation of symptoms, diagnosis of SCLC is confirmed by histological evaluation of bronchoscopy samples and study of fine-needle aspiration,

endoscopic ultrasound guided fine-needle aspiration or transbronchoscopic needle aspiration [11-14]. Utilization of these methods lack certain pathological discrimination of SCLC from other lung malignancies [12] and sampling errors are the most common cause of false negatives in FNA cytology [12, 15]. Since it is difficult to diagnose SCLC accurately, development of a method for SCLC diagnosis based on molecular biomarkers is crucial and this method can increase the efficacy of SCLC diagnosis.

1.1.3 Potential of biomarkers for lung cancer diagnosis / early diagnosis

Biomarkers are valuable tools to be utilized for diagnosis, however to be useful in the clinic biomarkers should fit some criteria: The biomarkers should (1) be quantifiable and reproducible, (2) have good PPV and NPV, (3) be measurable in suitable material, in small amounts and with little effort, (4) indicate a disease state, (5) have validated clinical use, (6) be accepted by the market, (7) be economical; and (8) be reimbursed by insurance companies [16]. There can be biomarkers for the risk detection for lung cancer as well as biomarkers for distinguishing malignant from benign tumors. Identifying individuals at high risk may prevent the false-positive rates and decrease the overall number of CT-scans to be performed both for the benefit of cost reduction and accurate diagnosis. Similarly a biomarker that can distinguish a malignant tumor from a benign tumor would be invaluable since it can reduce the number of unnecessary surgical operations which is applied to 24% of indeterminate pulmonary nodules and 30% of lung cancer patients.

1.2 IMMUNE SYSTEM AND CANCER

1.2.1 Immune response against cancer

Immune system and cancer interaction and interplay is valuable in determining the type of biomarkers to be utilized for cancer diagnosis. There is an interplay between cancer and immune system during cancer development which is called Immunosurveillance which both eliminates and manipulates the development of cancer. Immunosuppressed individuals have higher incidences of cancer subtypes compared to immunocompetent ones [17]. Immunosurveillance leads to the understanding of immunoselection; which constitutes the stages of elimination, equilibrium and escape by which cancer is manipulated to evolve by the immune system. Immunosubversion also contributes to a favorable growth of the cancer by suppressing the immune system [18]. So that immune system plays a major role during the

evolution of the tumor and there is a continuous immune response against cancer throughout its development. Both innate and humoral immunity takes role in this active response.

1.2.2 Serum reactivity against cell surface antigens

Some of the early studies about cancer immunity addresses the immune responses against the tumor cell surface antigens. These studies utilized hemadsorption and immune adherence assays to show serum reactivity profiles with surface tumor antigens. Absorption tests with cells from autologous, allogeneic and xenogeneic sources, determine the specificity of reactions between melanoma cells and autologous serum. By these studies it was shown that there is a heterogeneity of serum reactivity against surface antigens in melanoma. Patients have different serum reactivity profiles in terms of the tumor cell surface antigen identity [19].

1.2.3 Serum reactivity against intracellular antigens

Serum reactivity was not only observed against cell surface antigens but also against intracellular antigens. p53, one of the most famous tumor suppressor genes was originally identified as a transformation related antigen in induced sarcomas as a tumor associated antigen. p53 is a transformation related antigen consistently identified in methylcholanthrene induced sarcomas, MuSV and SV40 transformed cell lines, spontaneously transformed fibroblasts and leukemias, but not in nontransformed cells. Either induced or not, p53 autologous antibodies were consistently present in the sera of individuals that harbor a tumor transformation. Autologous serum from Meth A mice were shown to immunoprecipitate p53 from Meth A sarcoma extracts [20]. So that serum reactivity can also develop against intracellular antigens.

1.2.4 Methylcholanthrene induction causes tumor cells to be rejected by the host

Mutagenic agents like methylcholanthrene were also shown to cause the tumor cells to be attacked by the immune system. Methylcholanthrene induction of a nonimmunogenic malignant mouse teratocarcinoma resulted in immunogenic variant cells that were incapable of forming tumors in syngenic mice. These variants undergo a process of immunerejection and confer a long lasting immunity. The inability of the induced cells to form tumors was specifically due to the immune response since they were able to survive in the gamma

irradiated syngenic mice. Methylcholanthrene induction makes the non-immunogenic tumor cells express strong transplantation antigens which leads tumor rejection by the host [21].

1.2.5 Immunoediting is a driving force for less immunogenic tumor development

Immune system is known to be continuously shaping the tumor throughout its development. Immunoediting of tumor is a driving force for less immunogenic tumor development. Tumors developed in normal BALB/c mice are less immunogenic compared to tumors developed in T-cell deficient nude mice. The number of tumors that were not rejected by the normal mice were 2 fold of the ones that were originating from nude mice. This indicates that tumors are shaped by immunoediting toward a less immunogenic phenotype in time. Chemically induced sarcomas from nude mice are more immunogenic than similar sarcomas from congenic normal mice [22].

1.2.6 Autologous CTLs lyse antigen presenting tumor cells

Immune system is not only involved in less immunogenic tumor development, but also well known to actively kill tumor cells by a CTL response. MAGE-1 is a cancer-testis (CT) antigen against which a CTL response is developed through tumorigenesis. Autologous CTLs specifically recognize an immunodominant epitope of the protein which is a nonapeptide. The presentation of the nonapeptide on HLA molecules causes lysis of the tumor cells by the CTLs. The nonapeptide has superior effect on the lysis capacity of CTLs compared to the octapeptides lacking the N- or C- terminal amino acids thus the CTL response is specifically targeting the nonapeptide MHC complexes. MAGE-1 specific CTLs lyse tumor cells which present MAGE-1 nonapeptide on MHC Class I, but not the octapeptides lacking the N- or C-terminal amino acids [23].

1.2.7 Autologous antibodies as part of the immune response to cancer

Together with CTLs autologous anti-tumor antibodies are part of the humoral immune system that has the potential to be efficiently utilized in clinic as an indicator of cancer if profiled clearly. This is because autoantibodies are valuable since they are more stable and persistent in serum and so they can be measured by methods like ELISA compared to autoantigens causing their elucidation which are transient in duration due to short lived

changes in tumor site so requires a biopsy to monitor. However the antibody profiling is complex since tumor development is an evolutionary process under the influence of the immune system and autoantibody repertoire in the serum is highly probable to be patient specific and so it is hard to establish a general evaluation methodology to apply to all patients [18].

Autoantibody biomarkers have advantages over TAAs since they are stable in serum, can be present in high concentrations compared to a low level expression of the TAA which elicits the antibody, and a panel of the autoantibodies can have increased sensitivity and specificity values for cancer diagnosis [17]. Moreover these autoantibodies can also be utilized for the prognosis and immunotherapy of cancer as well as defining personalized therapy regimes [17]. Early development of autoantibodies implies their role in the elimination step of immunosurveillance. So that it is important to investigate the autoantibody elucidation mechanisms and reasons parallel to tumor progression for the sake of understanding the biology of tumorigenesis.

1.2.8 Autologous antibody elucidation mechanisms

Elucidation of a humoral immune response against tumor is not a well-established phenomena. However autoimmunity against tumor antigens is generated based on 4 known possible mechanisms: 1) a missense mutation in ORF of a gene causing a new epitope to be presented by the MHC class I for recognition as a foreign epitope by the T-cells, 2) a mutation in ORF of a gene causing a new epitope for recognition by TCR, 3) overexpression of a self protein, 4) ectopic expression of a self protein. Even intracellular proteins can become immunogenic by these mechanisms by usual immuno-monitoring mechanisms, as well as aberrant tumor cell death and consequent release of intracellular proteins into an immunogenic tumor environment and tumor cell microvesicle shedding [17].

Most of the autologous antibodies are elicited against intracellular proteins. Detailed mechanism of this phenomenon is not well understood. The current hypothesis is; the major component of the immune response against intracellular proteins is a CTL response and the B cell response is a side effect of this mechanism [24]. The tumor cells destructed in the tumor microenvironment either by mechanical force, phagocytosis, autophagy, apoptosis or necrosis result in recognition of the tumor specific antigens by the immune system. These antigens are primarily presented to CTLs and CD4⁺ T-cells and antibody elucidation develops as a secondary response.

1.2.9 Antigen expression and autologous antibody elucidation

Previously we have identified 2 proteins; SOX2 and ZIC2, against which high titer autologous antibodies are elicited in SCLC. SOX proteins are generally expressed within the developing nervous system of the vertebrate embryo, and their expression either ceases or is reduced at later stages of development. It was shown that in both chicken and mouse, Sox2 is expressed predominantly in the developing central nervous system and are down-regulated as neural differentiation progresses. Outside the nervous system, low-level mouse Sox group B gene expression has been reported in gut epithelium, developing limb buds, and genital ridges of the embryo. Mouse Zic genes, although unrelated to Sox group B genes, have similar temporal expression patterns, and Sox2 and Zic1 are subject to common developmentally regulatory elements. All Zic genes are expressed from the gastrula stage on and show different temporal expression patterns. Xenopus Zic2 expression commences earlier than XZic1 and 3. In the adult mouse, Zic2 expression is restricted to the granule cells of the cerebellum. Although SOX2 mRNA can be detected in some adult tissues, ZIC2 is expressed only in brain and testis of humans. Most probably various gene activation or derepression mechanisms are involved in the tumor expression of ZIC2 and maybe SOX genes. A likely explanation of why SOX2 and ZIC2 antibodies are elicited is that SCLC expresses a range of neuroectodermal antigens, including SOX group B and ZIC2, primarily restricted to the nervous system and that immune tolerance has not been established against these antigens. Growing outside the brain–blood barrier, SCLC would provide a potent immunogenic stimulus to the host. Of the SCLC cell lines tested in the study 80% were expressing ZIC2 mRNA whereas 50% were expressing SOX2 mRNA. SCLC patients tested in this study were 41% seropositive against SOX2 and 29% seropositive against ZIC2. This study showed that high-titered ($1:10^6$) IgG response to tumor antigens is not a rare event and such responses are not limited to patients with Paraneoplastic Neurological Disease (PND). None of the patients having SOX2 or ZIC2 antibodies have PND in the patient population. However the correlation between autologous antibody elucidation and antigen expression was not clarified with this study. So further study needs to be performed to evaluate the potential of these biomarkers as tumor biomarkers and the mechanisms triggering antibody elucidation [25].

1.2.10 Autologous antibodies and PNS

Autologous antibodies against tumor-associated antigens (TAA) is part of this immune response [26-32] and has the potential for early tumor diagnosis since it is well-known that

these antibodies are elicited even when the tumor is at its microscopic state [29, 33-35]. The autologous anti-tumor antibodies correlate with PNS in a tiny percent of cancer patients and these patients can be diagnosed with a cancer even after 10 years of PNS emergence [36]. In the case of SCLC-associated PNS the neurological disease is known to be symptomatic before the associated SCLC is symptomatic and thus diagnosed [37, 38].

There are a number of known autologous anti-tumor antibodies in SCLC. Most of these antibodies are indicators of presence of a PNS. About 6% of the SCLC patients are reported to have PNS (22); 3-4% of SCLC develop LEMS [39, 40] which is associated with VGCC autoantibodies. However SCLC patients without PNS are also shown to elicit autologous anti-tumor antibodies. For example SOX2, p53, NY-ESO-1, CAGE, GBU 4-5, Annexin 1 are shown to have sensitivity values of 11, 13, 13, 12, 9, 9 respectively, in SCLC patients without PNS [26].

1.2.11 Autologous antibodies in SCLC with/without PND

Autologous anti-tumor antibodies were observed in many cancers. Lung cancer is among these cancers and the autologous anti-tumor antibodies are not only elicited against tumor associated antigens (TAA) but also in the case of PND high titers of autologous antibodies are frequently observed. Table 1 summarizes the autologous antibody frequencies in cancer with and without PND.

Table 1: Autologous antibodies in cancer and PND

	Antibody	Cloned genes	Tumor type	Paraneoplastic disease	Frequency in SCLC w/o PND	Frequency in SCLC with PND
1	Anti-Hu (ANNA-1)	<i>HuD, HuC, Hel-N1</i>	SCLC , Neuroblastoma, Prostate	Encephalomyelitis, sensory neuropathy, cerebellar degeneration, autonomic dysfunction	18.5-25.5 % ^[41]	75 % ^[42]
2	Anti-ZIC	<i>ZIC4 (ZIC1,3?)</i>	SCLC	Cerebellar degeneration	16 % ^[42]	29 % ^[42]
3	Anti-Recoverin	<i>Recoverin</i>	SCLC , melanoma, gynecological tumors	Retinal degeneration	15 % ^[43]	66 %
4	Anti-CRMP5	<i>CRMP5 (POP66)</i>	SCLC , thymoma	Encephalomyelitis, sensory neuropathy, cerebellar degeneration, corea	9.5 % ^[42]	27% ^[42]
5	Anti-VGCC	<i>P/Q type VGCC</i>	SCLC	Lambert-Eaton myasthenic syndrome	5 % ^[41] -18 % ^[44]	91 % ^[44]

6	Anti-VGCC	<i>N type VGCC</i>	SCLC	Lambert-Eaton myasthenic syndrome	12 % ^[45] - 22 % ^[44]	73 % ^[44]
7	Anti-Amphyphysin	<i>Amphiphysin</i>	SCLC , Breast	Encephalomyelitis, stiff-man syndrome	1.4 % ^[46]	2.9 % ^[46]
8	Anti-PCA-2	?	SCLC	Cerebellar degeneration	?	
9	Anti-VGKC	<i>Potassium channels</i>	SCLC , thymoma	Neuromyotonia	?	
10	Anti-Enolase	<i>Enolase</i>	SCLC , breast, pancreas	Retinal degeneration	?	29 % ^[47]
11	Anti-p35	<i>p35</i>	SCLC	Retinal degeneration	?	5 % ^[47]
12	Anti-Yo (PCA-1)	<i>CDR34, CDR62</i>	Lung , ovarian, breast	Cerebellar degeneration	3 % ^[45]	
13	Anti-Ri (ANNA-2)	<i>Nova1, 2</i>	Lung , breast, ovarian, bladder, gynecological	Ataxia with or without opsoclonus-myoclonus	3 % ^[45]	
14	Anti-Ma1	<i>Ma1</i>	Lung	Brain-stem encephalitis, cerebellar degeneration	?	
15	ANNA-3	?	Lung	Sensory neuropathy, encephalomyelitis	?	
16	Anti-Tr	<i>MAZ</i>	Hodgkin's lymphoma, brain, lung , testis	Cerebellar degeneration	16 % (non-SCLC tumors) ^[48]	
17	Anti-Ma2	<i>Ma2</i>	Testicular	Limbic brain-stem encephalitis	?	
18	Anti-MAG	<i>MAG</i>	Waldenstrom's macroglobulinemia	Peripheral neuropathy	?	
19	Anti-mGluR1	<i>Glu receptor</i>	Hodgkin's lymphoma	Cerebellar degeneration		

The autologous antibodies associated with PND are commonly elicited in lung cancer; specifically in SCLC. So that SCLC is an immunogenic cancer. The frequency of autologous antibodies in SCLC with PND is always higher than the frequency of autologous antibodies in SCLC without PND. However SCLC patients without PND also elicit antibodies and this data suggests that there might be other autologous antibodies which are independent of PND. Thus the autologous antibodies that are specific to SCLC without PND can be exploited for diagnosis/early diagnosis of SCLC.

1.3 DIAGNOSIS AND EARLY DIAGNOSIS OF CANCER

1.3.1 Autologous antibodies precedes PND and clinical cancer diagnosis

The timing of elucidation of autologous antibodies is critical for their utilization for early diagnosis, thus their presence in the early stages of tumor development defines their potential. PND coexists with a tumor for a significant number of SCLC patients. Anti-Hu

antibodies are elicited in almost 70% of the SCLC patients having a coexisting PND. The symptoms of PND is observable in average of 6.5 months before cancer and PND diagnosis. Data strongly suggests that autologous anti-tumor antibodies are detectable at least 6.5 months before SCLC diagnosis [49]. Not only anti-Hu antibodies but several other antibodies are also detectable up to 5-10 years before cancer and PND diagnosis [49-51]

1.3.2 p53 antibody level stability / fluctuations can predict cancer development

Not only antibodies associated with PND precede clinical cancer diagnosis, but also antibodies independent of PND; p53 antibodies were shown to precede cancer diagnosis. Table 2 shows the p53 antibody levels in a time course of up to 12 years in case of cancer development or controls.

Table 2: p53 antibody levels in case of cancer development and controls

Cancer

Sample #		Disease	Anti- p53 antibody seropositivity (years)					Cancer	Method	Ref.
			≥ 4	-3	-2	-1	months			
1	PT37				18.3 months	+	+	SCC*	ELISA	[52]
2	24	VC	11(>10)					AS	ELISA	[53]
3	23	VC					4	AS	ELISA	[53]
4	12	COPD					7	lung	ELISA/IP/W#	[35]
5	10	COPD		(-)	(-)	(-)	6	lung	ELISA/IP/W	[35]
6	8	COPD				+	5	göğüs	ELISA/IP/W	[35]
7	11	COPD		+	+	+		prostate	ELISA/IP/W	[35]
8	39	BM					2	SCC-E	ELISA/IP/W	[54]
9	2	BM					1	AC-E	IP/W	[54]
10	1	ASB	+	+	+	+		lung	ELISA	[34]
11	2	ASB		+	+	+		lung	ELISA	[34]
12	3	ASB	6	+	+	+		lenfoma	ELISA	[34]
13	4	ASB	+		+			lung	ELISA	[34]
14	5	ASB			+	+		Mesoth.	ELISA	[34]
15	6	ASB	(-)	(-)	+			lung	ELISA	[34]
16	7	ASB			(-)	+		lung	ELISA	[34]
17	8	ASB	10					lung	ELISA	[34]
18	9	ASB			+			lung	ELISA	[34]
19	12	ASB				+		lung	ELISA	[34]
20	13	ASB	12					lung	ELISA	[34]

Healthy

Sample #	Patient No.	Disease	Anti- p53 antibody seropositivity					Method	Ref.
			1.year	2.year	3.year	4.year	5.year		
1	14	ASB*	+	(-)	(-)	(-)	(-)	ELISA	[34]
2	15	ASB		(-)	(-)	+	(-)	ELISA	[34]
3	16	ASB			+		(-)	ELISA	[34]
4	17	ASB		(-)	(-)	+	(-)	ELISA	[34]

These results indicate that anti-p53 autologous antibodies' presence can precede cancer diagnosis even up to 12 years. Also p53 antibody levels are stable if there is an underlying cancer development. However p53 antibody response is sporadic in persons without cancer and the antibody levels fluctuate.

1.3.3 Sensitivity and specificity of autologous antibodies

Even though autologous anti-tumor antibodies can be utilized for early diagnosis, the sensitivity values of single antibodies cannot exceed 15% at 97% specificity. Published data strongly suggests that these autologous anti-tumor antibodies are not limited to a group of cancer patients so that their utilization in early cancer diagnosis has a high potential [55]. It is strongly suggested that the evaluation of autologous antibodies as a panel can increase the sensitivity at a certain specificity for early cancer diagnosis.

1.3.4 Autologous antibodies already in clinical use for diagnosis of cancer

There is already a diagnostic kit for lung cancer; Oncimmune's EarlyCDT-Lung. The kit has a Sensitivity value of 40% at a Specificity value of 93% for lung cancer. SCLC has a relatively higher sensitivity compared to NSCLC, which also suggests that SCLC is a more immunogenic tumor type. There is no significant difference in terms of sensitivity of the kit between late stage and early stage of lung cancer or between limited and extended disease categories of SCLC or between NSCLC Stages I, II, III and IV [56] This is a strong indication that autologous anti-tumor antibodies can be utilized for diagnosis of any type of lung cancer independent from stage.

1.4 AUTOLOGOUS ANTIBODY DISCOVERY METHODS

1.4.1 Methods for autologous antibody discovery

Since autologous antibodies are valuable tumor markers new autologous antibodies need to be identified. There are a number of methods for novel autoantibody discovery. These

methods include serological analysis of expression cDNA libraries (SEREX), serological proteome analysis (SERPA), multiple affinity protein profiling (MAPPING), reverse-capture microarray, phage display, protein micro- and macro-array screening [17]. Each of these methods have their own specific limitations. By SEREX CT antigens (NY-ESO-1, SSX2, MAGE, etc), mutational antigens (p53, etc), differentiation antigens (tyrosinase, SOX2, ZIC2, etc) and embryonic antigens have been identified so far in cancers as immunogenic [25, 57, 58]. However the antigens identified by SEREX have low sensitivity values [17] and SEREX method; even though is one of the most powerful methods, has some limitations in identifying autoantigens [59]. PA (PA) screening has the potential for identification of many immunogenic clones in a specific cancer in a comparative analysis with the samples of benign form of the disease or healthy controls and in most cases further validation and analysis of these clones is possible by subarraying [60].

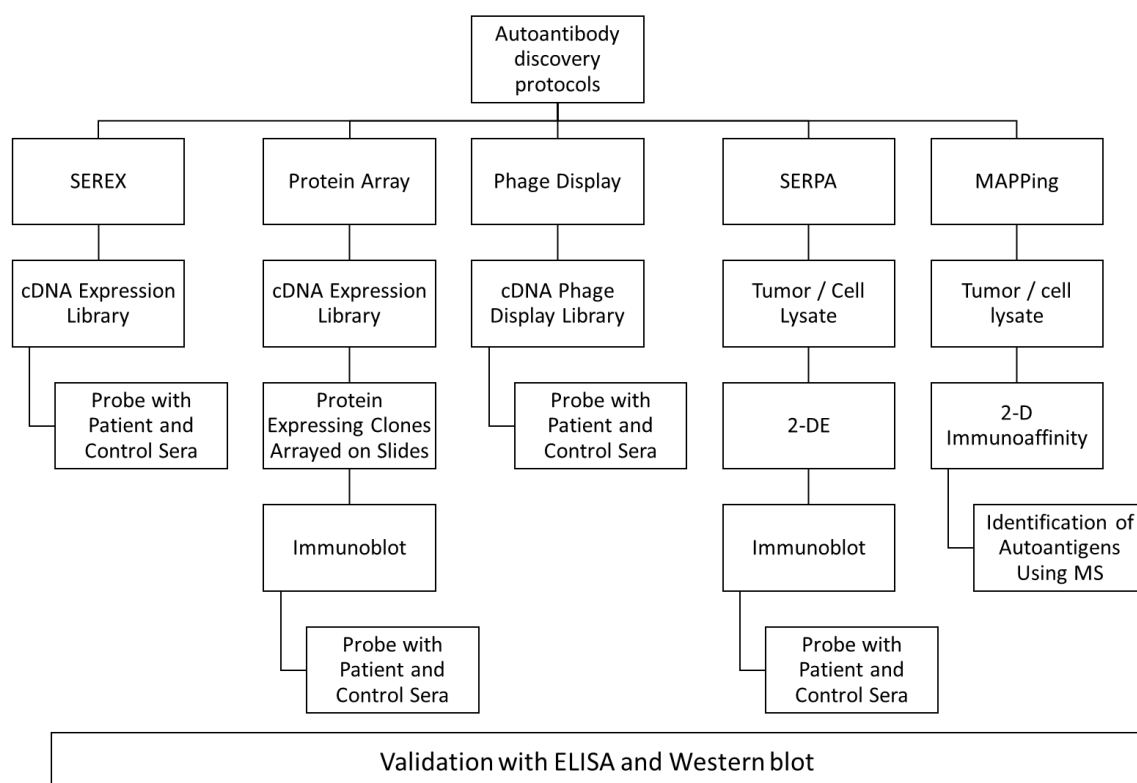


Figure 1. Conventional techniques for novel autologous antibody discovery and validation. As common features all techniques include a cDNA expression library or tumor / cell lysate screening and immunoblotting with patient and control sera.

Figure 1 summarizes the procedures of different available methods for identification of new autologous anti-tumor antibodies. The methods start with a cDNA library or a tumor/cell lysate and investigate the affinity of autologous anti-tumor antibodies to the starting material. To identify tumor specific antibodies cancer and healthy sera are screened in parallel and compared afterwards. Even though all these methods have differences in the way of their application at the end validation of the identified autologous anti-tumor antibodies with ELISA or Western blot is common to all methods. Otherwise it is not possible for these newly diagnosed antibodies to be utilized in clinic.

Table 3: Correlation between SOX2 phage and ELISA results

	SOX2 Phage	SOX1 Phage
Number of XY Pairs	90	90
Pearson r	0.6855	0.7181
95% confidence interval	0.5575 to 0.7817	0.6003 to 0.8055
P value (two-tailed)	< 0.0001	< 0.0001
P value summary	***	***
Is the correlation significant? (alpha=0.05)	Yes	Yes
R square	0.4699	0.5157

Previously we have investigated the correlation of our SEREX and ELISA results (Table 3). There was a significant correlation and this indicates if more valuable autologous anti-tumor antibodies can be identified by SEREX these will most probably correlate with ELISA.

Previously many autologous anti-tumor antibodies have been identified by SEREX and they are stored in a database named; Cancer Immunome Database. Even though more than 1500 cancer specific autologous anti-tumor antibodies have been identified till now most of them have a limited sensitivity and specificity. Because of this reason new methodologies need to be investigated to identify more valuable autologous anti-tumor antibodies.

A number of methods have been developed to identify the antigens against which the autologous antibodies are elicited. SEREX is one of the most powerful methods to identify novel tumor antigens. SEREX was performed for many cancers like colon, lung, melanoma, renal and Hodgkin's disease. The antigens identified by SEREX can be categorized into many different categories like mutational, differentiation, amplified or overexpressed, retroviral, splice variants and CT antigens. The CT antigens are mostly located on the X chromosome and almost all of them represent a family of genes [61].

One of the first autologous antibodies identified by our group is NY-ESO-1, which is a CT antigen. It was identified by utilization of SEREX in esophageal squamous cell carcinoma patients. NY-ESO-1 is a CT antigen, expressed mainly in testis, uterus and a diverse range of tumors. Northern blot results indicate NY-ESO-1 expression in Testis and melanoma cell lines. NY-ESO-1 is not expressed by the normal tissues. NY-ESO-1 has limited expression in Testis and cancer cell lines, but not in the other healthy tissues. [62] CT antigens behave very similar to NY-ESO-1 in terms of expression regulation. They were defined as ideal targets for immunotherapy because of their limited expression to tumors and the above mentioned tissues together with eliciting autologous anti-tumor antibodies in cancer patients.

1.4.2 Limitations of SEREX

Even though SEREX was widely used for identification of new autologous anti-tumor antibodies it has dramatic limitations. For example cDNA libraries utilized in SEREX obtained from lung cancer cell lines represent different number of seropositive clones when these libraries were screened with the same serum (Table 4). Besides the same cell line represents different number of seropositive clones when screened with 3 different serum pools. This indicates that a single cell line has limited representation of mRNA diversity and likewise a serum pool has a limited representation of antibody diversity in SEREX.

Table 4: Technical limitations of SEREX (A)

		Cell Lines			p (Chi.sq.)
		SK-LC-17	SK-LC-14	SK-LC-7	
SERA	Pool I (HI: Lu 192, 224, 230, 232, 233)	33 / 600,000 (1 / 18,000)	15 / 560.000 (1 / 37,000)	NA	0.02
	Pool II (HII: Lu 14, 17, 19, 90, 146)	5 / 400,000 (1 / 80,000)	NA	2 / 80,000 (1 / 40,000)	0.39
	Pool III (HIII: Lu 21, 29, 41, 92, 106)	3 / 240,000 (1 / 80,000)	NA	0 / 24,000	NA
p (Chi.sq.)		< 0.001 (HI vs. HII) < 0.01 (HI vs. HIII) 1 (HII vs.HIII)	NA	0.4	

Table 5: Technical limitations of SEREX (B)

		SERA	
		Pool IV Anti-NY-ESO-1 Ab (+)	Pool V Anti SOX1, 2, 3 Ab(+)
Cell Lines	SK-LC-14 (NY-ESO-1+) (2.4 x 10⁴ pfu)	NY-ESO-1 (-)	<i>NA</i>
	NCI-H740 (SOX1,2,3+) (5.6x10⁵ pfu)	<i>NA</i>	SOX3 (-)

SK-LC-14 is a cell line known to express NY-ESO-1 mRNA by qPCR and Pool IV was shown to contain anti-NY-ESO-1 antibodies by ELISA, however no reactivity was observed with SEREX (Table 5). Similarly NCI-H740 is known to express SOX3 mRNA by qPCR and Pool V was shown to contain anti-SOX3 antibodies by ELISA, however no reactivity was observed with SEREX. Both NY-ESO-1 and SOX3 are previously shown to be expressed in mid-abundance (100-500 copies in 500.000 pfus) in these cell lines by qPCR, meaning that they are not even rare transcripts. So that SEREX has a limited mRNA representation capacity and the serum pools cannot be used more than once so that it is highly unlikely to detect many novel autologous antibodies by SEREX.

Table 6: mRNA representation limitations of SEREX

Genes	Screened pfu numbers Expected in 500,000:		Error rate: if 1:3 of all ORFs are correct:
	mRNA number	mRNA percent (%)	Number of proteins having the correct amino acid sequence
Efl α	4,399	0.88%	1,466
Cytochrom Cox 1	1,802	0.36%	601
Clone 190B1	1,670	0.33%	557
Tubulin β	1,361	0.27%	454
40S Riboprotein S6	1,203	0.24%	401
40S Riboprotein S4	658	0.13%	219
60S Riboprotein L4	559	0.11%	186
GAPDH	539	0.11%	180
Efl β	441	0.09%	147
Calmodulin	210	0.04%	70
HSP KD71	184	0.04%	61
HSP KD90	171	0.03%	57
TNFR	79	0.02%	26
Clone 244D14	53	0.01%	18
Clone 241F 17	13	0.00%	4

Table 6 shows the number of mRNA numbers for different genes in a representative SEREX cDNA library on the left panel. The middle panel shows the percentage of these numbers in 500.000 pfus. And the right panel indicates the number of proteins having the correct amino acid sequence in an ideal case since only one third of the mRNA copies can

result in expression of the true ORFs. Based on these numbers one can expect the lower half of these genes on the table cannot be represented by the plaques in SEREX which indicates a significant limitation of the technique.

1.4.3 Theoretical comparison of SEREX with PAs

Common antibody signature for all cancers can be identified if a large number of signals can be analyzed simultaneously as in the PA screening format [60]. Antibody signatures can discriminate cancer from healthy as well as benign diseased individuals of the same organ. These antibody signatures can pave the way to identify antigens that take place in general metabolic pathways and targeted by the immune system. PA screening can result in identification of both in-frame and non in-frame clones to be immunogenic. However the newly identified antibodies are required to be validated by independent assays which is the bottleneck in utilization of an antibody biomarker in clinical diagnosis [60]. Previously many studies have been performed to identify accuracy, sensitivity and specificity values to discriminate cancer from healthy controls by many different methods [60].

Table 7: Summary of limitations of SEREX

Method	SEREX	PA
Technique	Unnormalized mRNA library screening	Screening of 15,744 different human proteins
Ideal serum dilution	1:100	1:25,000
Serum amount required	5 ml / run	0.2 ml / run
Sensitivity	1	~50

As shown in Table 7, PA includes more than 15,000 different human proteins as opposed to an unnormalized cDNA library screening by SEREX. Serum can be diluted to 1:25,000 for PA screening compared to 1:100 diluted for SEREX which indicates a major advantage for PA screening in terms of required sample amount. 25 times less serum is enough for one PA screening compared to SEREX; 200 μ l for PA screening and 5 ml for SEREX. Finally if the sensitivity of SEREX is assumed to be 1 while PA screening is at least 50 times more sensitive.

1.5 AUTOLOGOUS ANTIBODIES AND PROGNOSIS

1.5.1 Autologous antibodies can be utilized for prognosis

Autologous anti-tumor antibody response is a component of the anti-tumor immune response. One of the oldest studies addressing these antibodies shows a nice correlation, even though it is not statistically significant, between antibody response and good overall survival for the SCLC patients. In this study a number of antigens were shown to be recognized by the autologous antibodies in the patient serum according to Western blot data. After this study a major aim of the research became the identification of the individual antigens recognized by the autologous anti-tumor antibodies. This study suggested that autologous antibody response can be useful as a prognostic factor for cancer outcome prediction [63].

1.5.2 Autologous antibody positive sera have cytotoxic effect on cancer cell line

Previously antibody positive sera were shown to have cytotoxic effect in vitro. Anti-Hu antibody positivity is not only associated with PND but also represents a strong anti-tumor immune response from the anti-cancer perspective. Anti-Hu antibody bearing sera are shown to have cytotoxic effect on NT2 cells. Whereas this effect is not observed with the tested normal serum. Anti-Hu negative SCLC sera were not as toxic as anti-Hu positive serum, and the difference is statistically significant. The cytotoxic effect was not caused by the anti-Hu antibodies themselves, since the serum was also toxic to Hu negative cell lines and IgG depleted serum was still cytotoxic to cell lines [64].

1.5.3 Autologous antibody positivity correlates with *in-vivo* cytotoxicity

Cytotoxic effect of the immune response coexisting with the autologous antibodies is also observable in vivo. A patient with anti-Yo positive paraneoplastic cerebellar degeneration (PCD) have lost all her Purkinje cells in the brain. However the Purkinje cells are clearly observable and morphologically normal in a cancer patient without neurologic involvement [65]. This indicates that the antibody seropositivity can be a measure of the cytotoxic immune response.

1.5.4 Autologous antibodies and survival

Different studies investigated the existence of a correlation between anti-Hu, anti-VGCC and anti-p53 antibodies and survival. No correlation was identified between anti-Hu and anti-VGCC antibodies and survival for the SCLC patients. However a correlation was identified between anti-p53 antibodies and worse overall survival for NSCLC. This correlation was also observable for Squamous cell carcinoma whereas there was no correlation for Adenocarcinoma [41, 66]. So that a clear correlation between autologous antibody response and survival could not be formed yet.

1.5.5 Autologous anti-tumor antibodies identified by our previous studies

As a group we have also identified 2 antigens against which autologous anti-tumor antibodies are elicited; anti-SOX1,2,3 and anti-ZIC2. SOX proteins have more than 90% homology in their DNA binding HMG box domain. Thus autologous antibodies which recognize any of the SOX 1, 2 or 3 more than >90% of the cases recognize the other two antigens. SOX2 and ZIC2 autologous antibodies are high affinity and are present in high titer in serum; these antibodies are even detectable by SEREX at a million times diluted concentration [25].

1.5.6 Anti-SOX1 autologous antibodies are specific to cancer not PND

Autologous anti-SOX1 antibodies correlate with the existence of SCLC in Lambert-Eaton Myasthenic Syndrome. The seropositivity rate in the Paraneoplastic LEMS is 64% whereas none of the idiopathic patients are seropositive against SOX1. This data indicates that autologous antibody response against SOX proteins is highly specific to SCLC [67].

1.5.7 SOX2 antibodies and prognosis of multiple myeloma (MM)

Anti-SOX immune response is not limited to SCLC most probably since SOX2 is a stem cell antigen and is responsible for self renewal of the stem cells. Monoclonal gammopathy of undetermined significance (MGUS) represents a precursor lesion to MM. Detection of anti-SOX2 T-cells predicts a favorable outcome in patients with asymptomatic plasmaproliferative disorders. The patients with anti-SOX2 T-cells also had a significantly lower likelihood of disease progression, with a 2-year progression-free survival rate of 100% versus 30% compared with patients lacking anti-SOX2 T-cells with a statistically significant p

value. Therefore, immunity to SOX2 predicts the clinical outcome in patients with asymptomatic plasmaproliferative disorders [68].

1.5.8 SOX2 and ZIC2 antibodies and survival in SCLC

SOX2 and ZIC2 antibodies are among the most frequent antibodies elicited in SCLC patients. Previous study shows, although not statistically significant there is a trend for better overall and progression free survival for seropositive patients against SOX2 and/or ZIC2 [69]. According to this study most SOX1, 2 and 3 autoantibodies in serum remained consistent even after 6 months from SCLC diagnosis. All patients seropositive for SOX1 had SOX2 antibodies, and most were reactive with SOX3. None of the anti-SOX1 antibody-negative patients gained seroreactivity over the period of study. Over the periods tested, antibody titers remained within 10-fold in all seropositive patients. This indicates autologous SOX2 anti-tumor antibodies are reliable and stable biomarkers through the course of SCLC.

1.6 SENSITIVITY AND SPECIFICITY OF PANELS

1.6.1 Diagnosis by evaluating a panel of antibodies

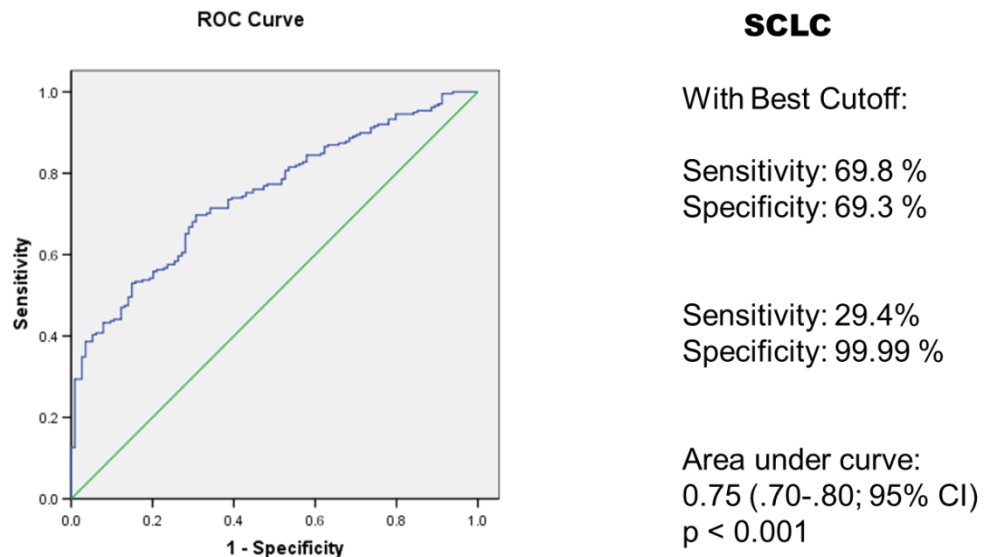
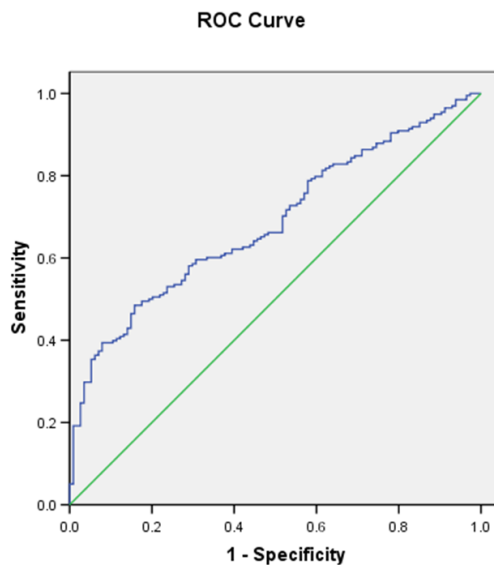


Figure 2. Anti-SOX2, HuD, CRMP5, p53, NY-ESO-1 and/or recoverin autologous antibodies in SCLC diagnosis. Green line indicates $x=y$, blue line is the ROC line. Sensitivity and specificity values were shown and AUC was indicated by confidence interval for $p < 0.001$.

Combined evaluation of 6 of the SEREX-identified autologous antibodies has 30% sensitivity at 100% specificity for SCLC diagnosis (Figure 2). If evaluated individually sensitivity of each antibody doesn't exceed 15%. This indicates that combined evaluation of more autologous antibodies thus forming a panel can increase the sensitivity for diagnosis even more. To increase the sensitivity a panel should be formed and new valuable autologous antibodies needs to be identified to be incorporated in the panel.



NSCLC

With Best Cutoff:

Sensitivity: 48.5 %
Specificity: 84.2 %

Sensitivity: 29.8%
Specificity: 96.5 %

Area under curve:
0.69 (.63-.75; 95% CI)
 $p < 0.001$

Figure 3. Anti-SOX2, HuD, CRMP5, p53, NY-ESO-1 and/or recoverin autologous antibodies in NSCLC diagnosis. Green line indicates $x=y$, blue line is the ROC line. Sensitivity and specificity values were shown and AUC was indicated by confidence interval for $p < 0.001$.

The situation is similar for NSCLC in which a sensitivity of 30% can be reached at 97% specificity with the same antigens evaluated for SCLC (Figure 3). So novel autologous anti-tumor antibodies needs to be identified to be utilized as a panel for cancer diagnosis

1.6.2 Sensitivity and Specificity of already identified autologous antibodies

The already identified autologous antibodies have significant sensitivity and specificity both for SCLC and NSCLC (Table 8). However these values values are not adequate for any of them or a panel of them to be utilized in clinical screenings for diagnosis.

Table 8: Sensitivity and specificity of anti-tumor antibodies in lung cancer

Sensitivity and Specificity of anti-tumor antibodies in Lung Cancer						
Positivity cut-off	p53			p53 + NY-ESO-1 + SOX2		
	Sensitivity		Specificity	Sensitivity		Specificity
	NSCLC	SCLC		NSCLC	SCLC	
M+3SD	20.7%	38.7%	94.7%	39.4%	54.6%	77.2%
M+8SD	7.1%	22.3%	99.1%	25.3%	36.6%	93.0%

Table 9: Sensitivity values of autologous antibodies useful in SCLC diagnosis at 100% specificity

Sensitivity values of autologous antibodies useful in SCLC diagnosis at 100% Specificity

Antigen	# of screened patients	Sensitivity when outliers are included (# and %)	Sensitivity when outliers are excluded (# and %)
NY-ESO-1	238	4 (%1.68)	30 (%12.61)
SOX2	238	38 (%15.97)	75 (%31.51)
p53	238	23 (%10.08)	24 (%10.08)
HuD	238	24 (%9.66)	36 (%15.13)
Combined seropositivity	238	75 (%31.93)	136 (%57.14)
Average seropositivity	238	22 (%9.35)	41 (%17.33)

For example p53 antibodies have 21% and 39% sensitivity for NSCLC and SCLC respectively. When combined with NY-ESO-1 and SOX2 the sensitivity values increase to 25 and decrease to 37% for SCLC and NSCLC respectively at a similar specificity. This indicates new autologous antibodies need to be identified to increase sensitivity without decreasing specificity. The sensitivity values of already identified autologous antibodies increases at 100% specificity if outliers are excluded for SCLC. For example SOX2 sensitivity increases from 16% to 32% (Table 9). The outliers can be excluded if antibodies are evaluated longitudinally for the same serum.

1.6.3 Predicted Sensitivity and Specificity of novel autologous antibodies

Table 10: Estimated effect of novel autologous antibodies in sensitivity of SCLC diagnosis

	Estimated sensitivity of novel autologous antibodies:					
	Increase rate	# of novel autologous antibodies				
		1	5	10	15	20
Sensitivity when outliers are included: %31.93	At least (%)	33.8	37.5	42.6	47.2	51.5
	Average (%)	38.3	58.3	74.5	84.4	90.4
	At most (%)	42.8	71.5	88.1	95	98
Sensitivity when outliers are excluded: %57.14	At least (%)	62.6	74.8	85.2	91.3	94.9
	Average (%)	64.6	83.5	93.6	97.5	99
	At most (%)	70.6	93.5	99	99.9	100

Table 11: Sensitivity and specificity values of autologous anti-tumor antibodies for other solid tumors

Tumor type / Antigens	n		M+2SD	M+3SD	M+5SD	M+10SD
Breast cancer / <i>NY-ESO-1, SOX2, HuD</i>	35	Sensitivity (%)	20	17.1	17.1	11.4
		Specificity (%)	88.4	90.7	95.1	100
Melanoma / <i>NY-ESO-1, SOX2, HuD</i>	43	Sensitivity (%)	18.6	18.6	18.6	2.3
		Specificity (%)	87.6	87.6	95.5	100
Ovarian cancer / <i>NY-ESO-1, SOX2, p53</i>	35	Sensitivity (%)	42.9	34.3	25.7	8.6
		Specificity (%)	94.2	92.9	96.7	100

The predicted effect of new autologous antibodies on SCLC diagnosis can be estimated if they are assumed to be SOX2 like in terms of sensitivity and specificity (Table 10). For example if there is only one autologous antibody its sensitivity can be predicted to be at least 34%, at most 43% and average 38%. If the number of new autologous antibodies increase to 10 the combined output of these values becomes 43%, 88% and 75% respectively. This scenario is in case the outliers are included. If the outliers are excluded each of these numbers increase to more than 90 for 10 antibodies.

The already identified autologous antibodies are not expected to be specific to SCLC only since the already identified antibodies have significant sensitivity and specificity values for breast cancer, melanoma and ovarian cancer (Table 11). For example combination of NY-ESO-1, SOX2 and p53 have more than 34% sensitivity at 93% specificity for ovarian cancer. This indicates new autologous antibodies specific to SCLC needs to be identified since it is highly probable that they will be valuable for diagnosis.

1.6.4 Commercially available diagnostic technologies

There is only one commercially available kit for lung cancer diagnosis which is composed of a panel of autologous antibodies. This kit is EarlyCDT-Lung and aims the combined evaluation of anti-p53, NY-ESO-1, CAGE, GBU4-5, Annexin 1 and SOX2 antibodies. The sensitivity of the kit is 40% at 90% specificity for lung cancer diagnosis. Individual sensitivity values of p53, NY-ESO-1 and SOX2 are 11%, 12%, 7% respectively, compared with the 35% panel sensitivity at >95% specificity. This data indicates that if more autoantibodies are included in the panel the sensitivity can increase [26].

Oncimmune increased the number of autologous antibodies to be evaluated in its EarlyCDT-Lung panel from 6 to 8. This increase resulted in a 2% increase both for sensitivity and specificity to 41% and 91% respectively. This indicates if the number of autologous antibodies in the panel are increased both sensitivity and specificity can be increased, and these novel autologous antibodies will contribute to the panel sensitivity and specificity proportional to their individual sensitivity and specificity values.

Chapter 2

2 AIM, RELEVANCE AND RATIONALE

2.1 AIMS

1) Development and optimization of a novel PA based method for autologous antibody discovery

- a) Serum selection for pooling
- b) Serum dilution ratio optimization
- c) Pool dilution ratio optimization
- d) Secondary antibody dilution ratio optimization
- e) Q-Dot dilution ratio optimization

2) Novel autologous antibody identification utilizing PAs and determining their diagnostic value

- a) Serum pool screening of PAs for SCLC, Gastric Cancer, Colorectal Cancer and Ovarian Cancer
- b) Individual serum screening of custom PAs for sensitivity and specificity calculations for SCLC, Gastric Cancer, Colorectal Cancer and Ovarian Cancer
- c) Identification of the most valuable clones for SCLC

3) Validation and expansion of PA data by ELISA and Western blotting

- a) ELISA screening of all sera with the most valuable antigens
- b) Cut-off determination
- c) Correlation analysis of PA, custom PA, ELISA and Western blotting data

4) Clinical correlates of autologous antibodies

- a) Univariate Cox Regression analysis for clinical parameters and seropositivity

- b) Multivariate Cox Regression analysis for clinical parameters and seropositivity
- c) Correlation analysis of seropositivity at multiple cut-offs with survival

5) Correlates of antigen expression, presence of autologous antibodies and clinical parameters

- a) Correlation analysis of SOX2 protein expression by IHC and seropositivity
- b) Correlation analysis of antigen expression intensity and frequency with survival

6) Evaluation of autologous antibodies in autoimmune disease

- a) Identification of autologous antibodies in NBD
- b) ELISA based validation of autologous antibody levels

2.2 RELEVANCE

The relevance of this effort is to determine the utility of autologous anti-tumor antibodies in the early diagnosis, differential diagnosis, prognosis prediction and their potential use as part of immunotherapy protocols in cancer.

2.3 RATIONALE

Our previous data suggests that there are many more antigens capable of eliciting autologous antibody responses in cancer, compared to what has been discovered thus far. Also, most of the previous studies have not included detailed sensitivity/specificity analyses, have not been tested in large cohorts and have not been checked for validity utilizing more than one antibody determination technique. As we possess serum from 2 large SCLC cohorts with known clinical outcome, we were able to test the clinical relevance of autologous antibodies, which has been pursued as part of this thesis. Another important question is why autologous antibodies are elicited in some patients and not in others. We have pursued an answer to this question by comparing antigen expression with seropositivity profiles in patients. Lastly, we contributed to the question of whether autologous antibodies could be utilized in autoimmune disease. The autologous antibodies that will be identified and validated for each specific disease have a potential to be utilized for early diagnosis of this disease, since autologous antibodies are elicited when the tumor is microscopic thus cannot be detected by conventional diagnostic methods. Further separate studies need to be performed for evaluation of the early diagnosis potential of each antibody.

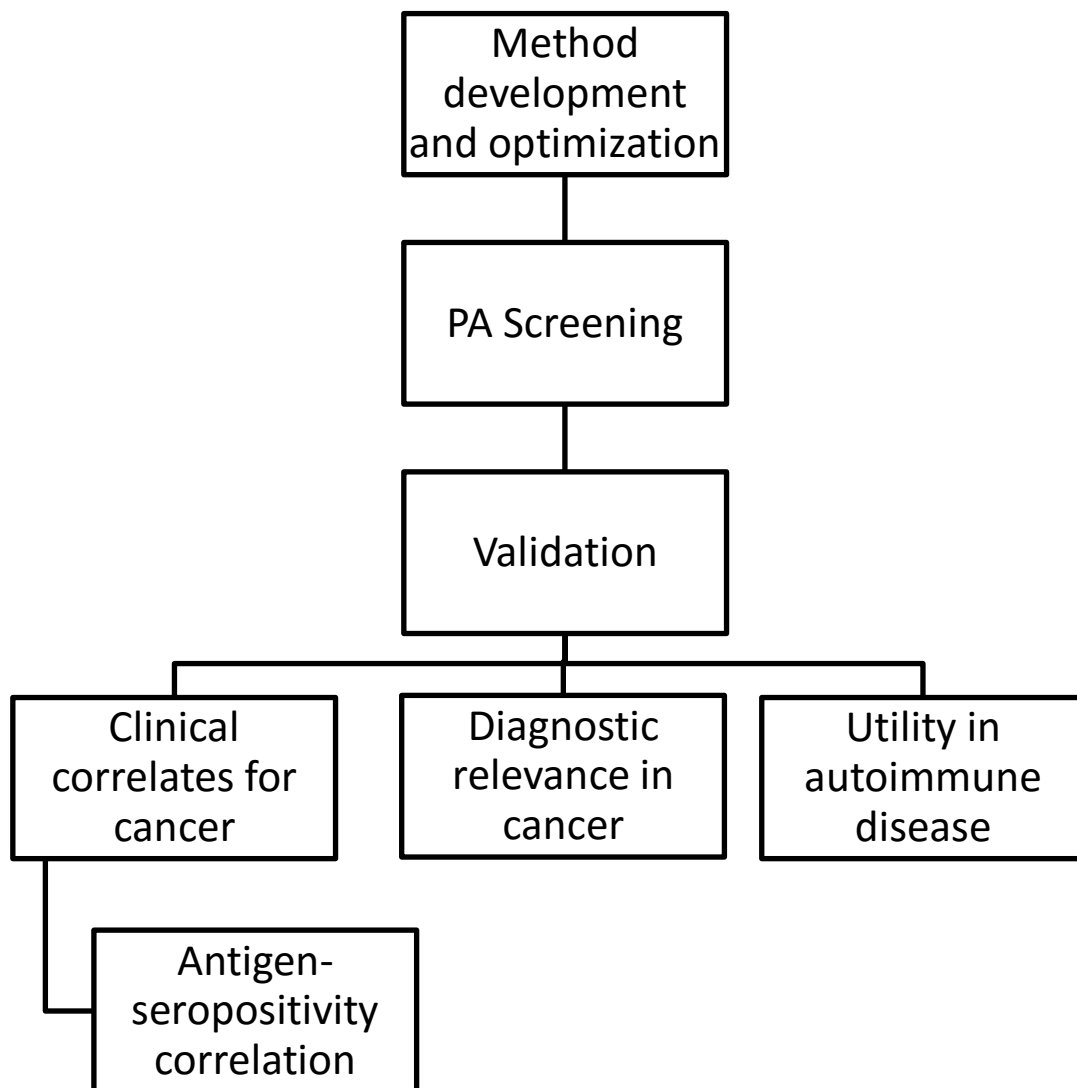
Each autologous antibody identified and validated in this study needs to be validated for its disease specificity by screening large cohorts of the disease, its benign forms and healthy controls.

The autologous antibodies identified in this study as correlating with survival needs to be validated by prospective studies and large cohorts to be utilized for prognosis prediction.

If the antigens against which the autologous antibodies identified as correlating with good prognosis in this study, have a limited expression profile to tumor tissue; they have the potential for utilization in immunotherapy protocols.

Chapter 3

3 STUDY DESIGN



Chapter 4

4 MATERIALS AND METHODS

4.1 METHOD DEVELOPMENT

4.1.1 Optimization Experiments for PA screening protocol determination

For optimizing the PA screening protocol we have performed optimization for the following steps:

- Serum selection for pooling
- Serum dilution
- Pool dilution
- 2^o Antibody dilution
- Q-Dot dilution

4.1.2 Serum Selection for Pooling

We have cut 0.5 x 2 cm PVDF membranes and these membranes were wetted in methanol for 30 seconds, in ddH₂O for 30 seconds and in TBS-T for 1 minute. Standard Western blotting protocol (explained in the Western blotting section) was applied to these membranes. Each stripe was treated with a single serum at the primary antibody incubation step. The membranes resulting intense background staining were selected and excluded from pooling.

4.1.3 Serum Dilution

For optimizing the ideal serum dilution in PA screening we have tried 3 different serum dilution ratios; 1:3000, 1:10000 and 1:100000. SOX2 protein was run on SDS-PAGE and a SOX2 seropositive serum was utilized as the primary antibody. Standard Western blotting protocol (explained in the Western blotting section) was applied, only the difference is the primary antibody (serum) dilution ratios for each protein lane.

4.1.4 Pool Dilution

For optimizing the ideal pool dilution in PA screening we have tried 3 different pool dilution ratios; 1:5000, 1:10000 and 1:25000. We planned to pool 50 sera so that the corresponding pool dilutions will be 1:100, 1:200 and 1:500. p53, HuD and SOX2 were spotted on PVDF membrane. p18 was also spotted as the positive control. Seropositive sera against p53, HuD and SOX2 were pooled to be utilized as the primary antibody. Standard immunoblotting protocol (explained in the Western blotting section except for the SDS-PAGE) was applied, only the difference is the primary antibody (serum) dilution ratios for each strip.

4.1.5 2° Antibody dilution

For optimizing the ideal 2° Antibody dilution in PA screening we have tried 2 different dilution ratios; 1:2000 and 1:500. SOX2 protein was run on SDS-PAGE and a SOX2 seropositive serum was utilized as the primary antibody. Standard Western blotting protocol (explained in the Western blotting section) was applied, only the difference is the 2° antibody dilution ratios for each protein lane.

4.1.6 Q-Dot dilution

For optimizing the ideal Q-Dot dilution in PA screening we have tried 3 different dilution ratios; 1:2000, 1:500 and 1:100. SOX2 protein was run on SDS-PAGE and a SOX2 seropositive serum was utilized as the primary antibody. Standard Western blotting protocol (explained in the Western blotting section) was applied, only the difference is the Q-Dot dilution ratios for each protein lane.

4.1.7 Strategy for autologous antibody discovery

As a strategy for identification of novel autologous antibodies we have planned to screen PAs with pooled sera. We have identified valuable clones capable of differentiating SCLC from healthy controls and rescreened those clones on custom PAs. We have screened these custom PAs with individual sera for sensitivity and specificity. Based on these values we have evaluated the clones and filtered the most valuable ones. These clones were utilized to produce pure proteins and these proteins were screened with ELISA and Western blotting for validation. Afterwards we have tried to correlate seropositivity with clinical correlates for all valuable proteins.

Figure 4 summarizes our experimental strategy both for identification and validation of novel autologous antibodies.

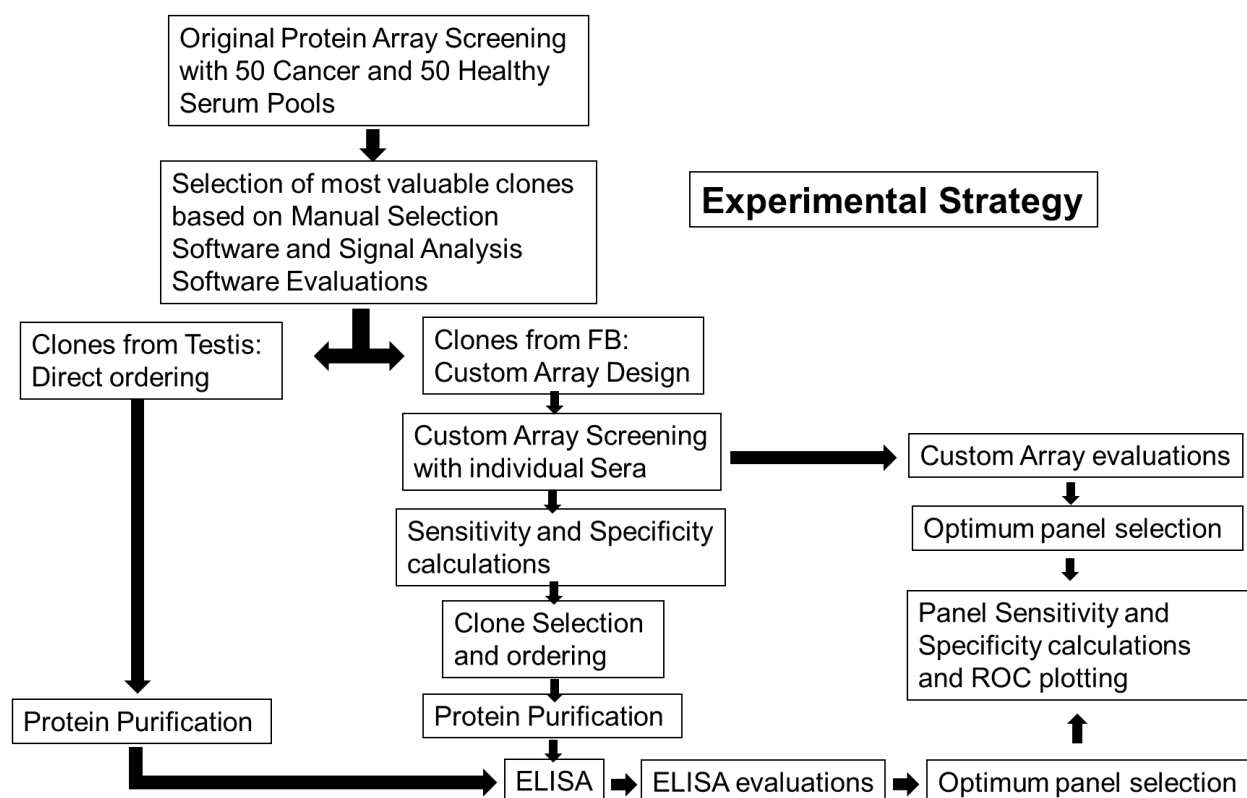


Figure 4. Complete experimental strategy for discovery and validation of novel autologous antibodies. The strategy includes pool screening, image analysis, custom PA screening, HisTagged protein expression and purification, ELISA, sensitivity and specificity calculations and panel formation.

First we have screened PAs with pooled sera in parallel both from 50 SCLC patients and 50 healthy controls. Based on these screening we have identified the most valuable clones that are capable of discriminating SCLC from healthy controls. We have performed this by utilizing 2 softwares; Manual Selection Software and Signal Analysis Software.

After this step we have applied 2 different strategies. For the clones selected from Testis screening we have directly ordered the clones from the PA developer company. For the clones selected from Fetal brain (FB) we have designed custom PAs by the most valuable clones selected from the pool screenings. We ordered these custom PAs from the PA developer company and screened each array with an individual serum. Based on these screenings we calculated sensitivity and specificity. Based on these calculations we have selected and ordered the most valuable clones.

We performed HisTagged affinity purification to the clones both from Testis and FB PA screenings and utilized the pure proteins in the following ELISA experiments. Based on these screenings we have evaluated the ELISA results for multiple SD cut-offs. Based on these evaluations and MC analysis we have selected an optimum panel of antigens. Besides we have also evaluated custom PA screenings by converting the signals to numeric values. Again we have selected an optimum panel based on these evaluations by utilizing MC algorithms. We have performed sensitivity and specificity calculations and ROC plotting for both of these panels and compared the results.

4.1.8 Protein Arrays

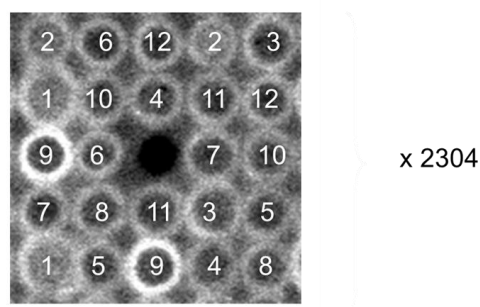


Figure 5. PA format and a representative positive signal. ImaGenes PAs have units of 25 spots; one reference InkDot (in the middle of the unit) and 12 clones spotted as duplicates. Each 22 cm x 22 cm PVDF membrane includes close to 27.000 spots (2304 units / membrane).

The PAs we have screened are two types: FB and Testis. These PAs were generated by transformation of FB and Testis cDNA libraries in bacterial expression vectors to obtain E.coli colonies. Clones were generated by cloning the cDNA libraries either in pQE30NST or pQE80LSN vectors, both of which are IPTG inducible. These arrays are spotted on 22 x 22 cm PVDF membranes. Testis PA consists of 26880 clones spotted as non-redundant on one 22 x 22 cm PVDF membrane representing about 15.000 distinct proteins whereas Feral Brain PA consists of 55296 clones spotted on 2 PVDF membranes the same size as Testis PA. Each PA is composed of units in each there are 12 different clones spotted in a duplicate fashion according to defined patterns. So each unit is composed of 25 spots; 2 x 12 clones = 24 spots + 1 InkDot (reference) = 25 spots. Each PA contains 2304 of these units (Figure 5). The spots on both PAs are E.coli colony lysates each of which inducibly expressed a specific cDNA clone from the specified library. The clones were lysed on the PVDF membrane so that each

spot contains the protein product of the cDNA clone together with all the bacterial proteins. (Figure 6) This characteristic of the PAs resulted in development of a comparative background signal from all clones together with true positive signals which led us in developing consistent and reliable evaluation softwares for the screening results.

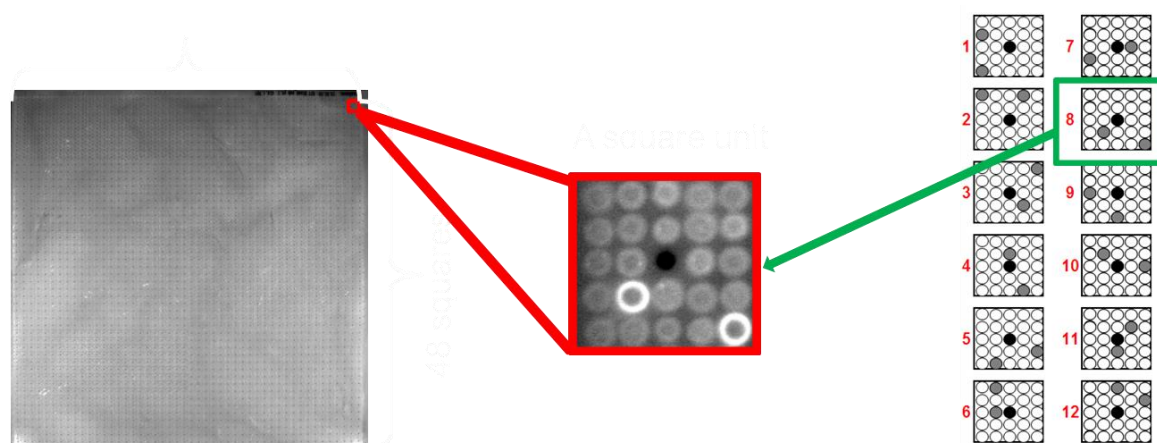


Figure 6. PA Spotting pattern is composed of 12 proteins in duplicate and a reference InkDot. The entire PA picture is shown (left), a positive signal harboring unit is magnified and shown in the box (middle), spotting patterns for 12 distinct clones is shown (right).

4.2 NOVEL AUTOLOGOUS ANTIBODY IDENTIFICATION

4.2.1 Screening Strategy and Cohorts

We had 3 screening strategies. These were screening the cohorts with pooled sera, screening with individual sera for validation and the screening for ELISA validation. The cohorts we have screened with pooled sera were SCLC, Ovarian, Colorectal, Gastric, BD and Healthy controls. The same cohorts were also screened for validation with individual sera. For ELISA validation we had 3 SCLC cohorts and 2 healthy cohorts.

4.2.2 Patient and Healthy Control Cohorts

PA screening was performed for SCLC, Colorectal, Gastric, Ovarian cancer patients, BD patients and Healthy controls. For each group 50 sera were selected. Patient sera were collected at the time of diagnosis before any treatment is applied. SCLC sera were collected at Ataturk Gogus Hastaliklari ve Gogus Cerrahisi, Arastirma ve Egitim Hastanesi, colorectal and gastric sera were collected at Ankara Numune Hastanesi, ovarian sera were collected at Ankara Zekai Tahir Burak Kadın Sagligi Egitim ve Arastirma Hastanesi, BD sera were

collected at Istanbul Universitesi Deneysel Tip Arastirma Enstitusu and healthy control sera were collected at Bilkent University Health Center. All patients gave informed consent and the study was approved by the ethical boards of the mentioned hospitals.

The patient population used for ELISA is composed of 3 independent SCLC cohorts and 2 healthy control cohorts. The first SCLC cohort is composed of 96 patients, the second 244 and the third 106 patients. The first Healthy control cohort is composed of 35 people and the second is composed of 157 people. SCLC cohorts 2 and 3 had clinical data available whereas cohort 1 did not. Serum was collected from all the patients at the time of diagnosis and none of the patients had neurological symptoms or evidence of PND within the follow-up period. All the patients had pathologically confirmed SCLC. All patients gave informed consent and the study was approved by the ethical board of the Ataturk Gogus Hastaliklari ve Gogus Cerrahisi, Arastirma ve Egitim Hastanesi of Ankara, Turkey and the ethical board of Capa Gogus Hastaliklari ve Gogus Cerrahisi, Arastirma Hastanesi of İstanbul, Turkey. All samples were anonymized before analysis. The clinical data obtained from patients at the time of diagnosis included age, gender, tumor stage, presence of metastasis, performance status (ECOG), serum alkaline phosphatase, and lactate dehydrogenase levels. All patients received chemotherapy with or without concurrent and sequential radiotherapy. Patients' response to therapy and type of progression (local/lung or distant) were recorded. Survival data was available for 300 patients. Follow-up times ranged from 0.2 to 44.7 months with a median of 8.68 months. Relapse was observed in 154 patients.

4.2.3 PA Screening

PAs were prepared for screenings according to the manufacturer's recommendations. WesternDot 625 Western Blotting kit (Invitrogen) was utilized throughout the entire screening procedure. PAs were agitated in blocking buffer for 1 hour at RT. Sera from 50 SCLC patients and 50 healthy controls were pooled separately and these pools were diluted 1:500 in wash buffer. One PA was agitated in each serum pool for 16 hours at RT. PAs were washed 3 times for 5 minutes in wash buffer with vigorous shaking. Goat anti-human (H+L Specific) ZyMAX Grade (Invitrogen) 2°Antibody is diluted 1:2000 in wash buffer and PAs were agitated in this solution for 1 hour at RT. PAs were washed 3 times for 5 minutes in wash buffer with vigorous shaking. Q-Dot Conjugate is diluted 1:1000 in blocking buffer and PAs were agitated in this solution for 1 hour at RT. PAs were washed 3 times for 5 minutes in wash buffer with vigorous shaking. PAs were washed in distilled water for an additional 5

minutes and left for drying on a Whatman paper at RT. PAs were visualized with StarLight module of Vilber Lourmat Gel Documentation System by utilizing the blue light and red filter of the apparatus. The optical zoom of the camera was adjusted to be maximum and all pictures were taken at this setting. A total of 48 photos were taken for each PA to cover the entire array at a sufficient resolution for reliable evaluation of the results.

4.2.4 Clone Selection Strategy based on PA Screening Results

PA screening results were first evaluated by Manual Selection Software (MSS) and all the signals were selected and manually sorted from the strongest to the weakest. The result is a mosaic image in which the background intensities of all signals were normalized by MSS and the signals became comparable by eye. Then the mosaic image is evaluated by the Signal Analysis Software (SAS) and all signals were transformed into numerical values. The signals giving an X value >5 in the mosaic image were selected for further evaluation.

Based on this selection strategy 27 clones from Testis PA were directly ordered and obtained. Imagenes provided custom PA design and ordering for the clones selected from FB PA for further evaluation. So we selected 40 clones from FB PA screening together with 140 clones from other cancer screenings for spotting on custom arrays. Custom arrays were ordered and obtained for individual serum screening of the selected clones.

4.2.5 Manual Selection Software (MSS) Development for selection of signals from PA screening results

MSS is a user dependent signal selection software which uses the 48 photos taken as a result of the PA screening. The user selects the PA units which contain a positive signal by clicking on the InkDot of the unit. Then the user types the pattern information of the selected signals. The software extracts the PA unit image from the rest of the photo and it also extracts exactly the same unit image from the parallel screening and puts the two unit pictures next to each other after performing a background normalization to make the units comparable. A representative output of MSS is shown in Figure 7. This procedure is applied to all selected signals and the software forms a mosaic picture representing all selected signals. After selection the software also performs signal sorting and deletion based on user decision. The resulting mosaic image is used by SAS to get numeric values from signals.

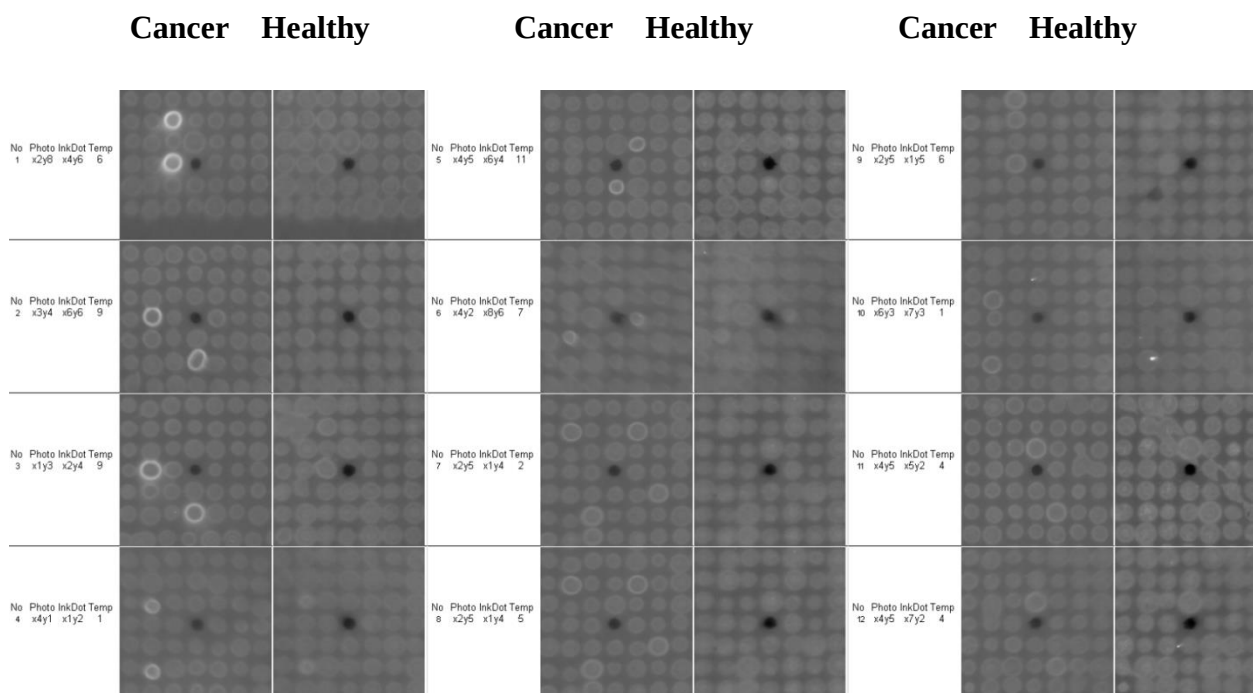


Figure 7. Representative output of MSS. Signals identified by visual evaluation are sorted from strongest to the weakest according to the signal intensity for cancer. Each unit harboring the signal is shown both for the cancer screening (left) and the healthy screening (right). The signal number, photo number, InkDot number and pattern number are shown on the left of each pair. Only the strongest 16 signals are shown due to lack of space.

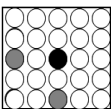
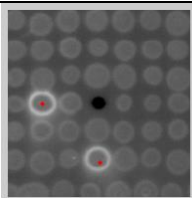
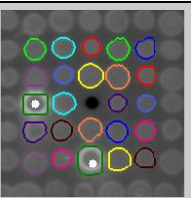
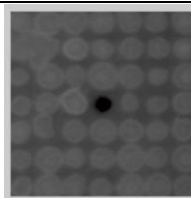
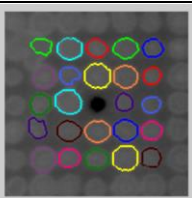
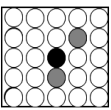
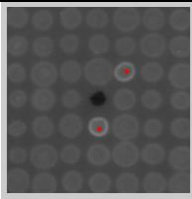
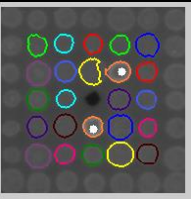
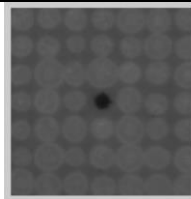
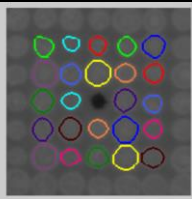
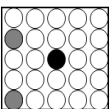
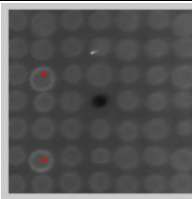
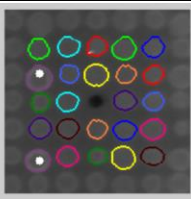
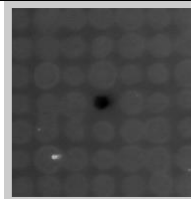
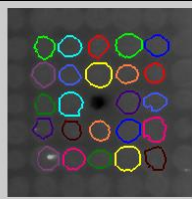
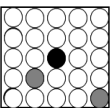
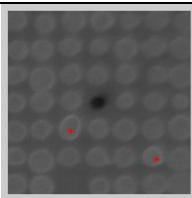
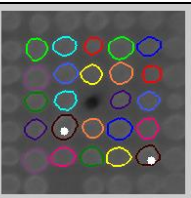
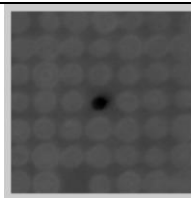
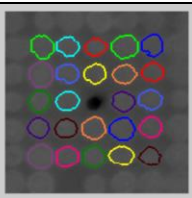
4.2.6 Singnal Analysis Software (SAS) and Pixel Count (PC) Method for Numerical Analysis

SAS is a software developed in MatLab mainly performing image processing on the PA units on mosaic photos generated as an output of the MSS. The software analyzes each PA unit independently and defines the border of each signal by utilizing adaptive threshold algorithms of MatLab. Since the software knows the spotting pattern information it automatically defines duplicate spots and represents their borders with identical colors. As a next step the software calculates the average of all pixels in the 4 pixel wide ring area to the inner side of the border of each signal which we call as the pixel value of each signal. Then the average (Mean) and standard deviation (StDev) of minimum 16 of all these 24 pixel values were calculated. Each signal was represented as the conversion of the pixel value to X values, which is represented in the formula: $\text{Mean} + X \cdot \text{StDev}$. Also the software is capable of indicating the positive signals based on a cut-off value. We have set 5SD and 15SD as cut-offs and performed all the analysis based on these values. A representative output of SAS is

shown in Table 12. The average of the X values for the duplicates were calculated for each of 12 clones and represented in an Excel sheet as an output. In the Excel sheet false positives were also eliminated by 2 different criteria:

If the difference between duplicate X values is more than 4 and if the smaller one is less than 2 or if the difference between duplicate X values is more than 10 and if the smaller one is less than 3 then the value of the smaller one is assigned to the bigger one.

Table 12: Representative output of SAS

No	Pattern	SCLC			Healthy		
		Signal	Borders	X*	Signal	Borders	X*
1				17.3			1.5
2				15.6			3.7
3				8.8			0.4
4				5.0			-0.3

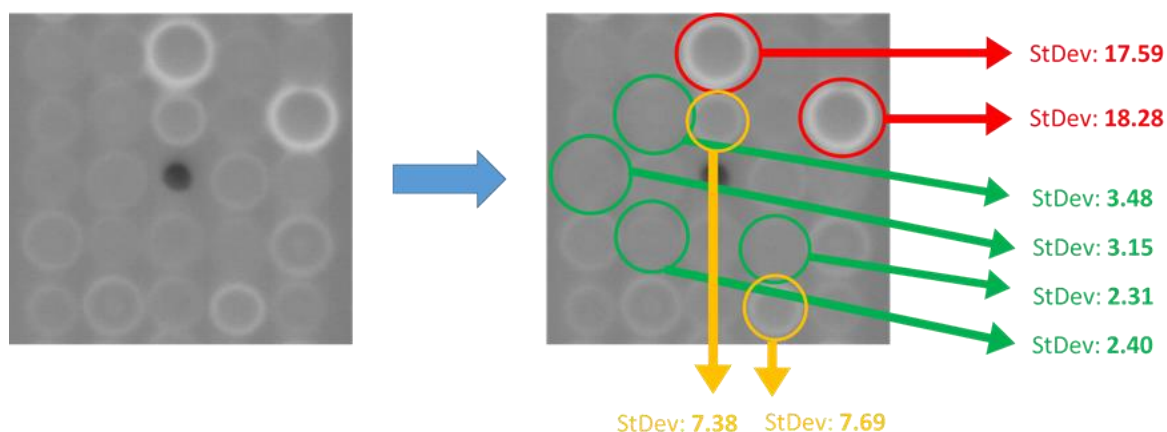


Figure 8. Evaluation method of PC. Each signal circle was manually selected tangentially to be completely included in the selection by Adobe Photoshop CS1. From the Histogram option the StDev value was obtained. This value was divided by a fixed StDev value of 2.5. The original unit containing strong, medium and weak signals (left) was evaluated by PC. The StDev values are indicated for representative signals.

We have also developed another quantitative analysis method that we call PC. In this method the pixels forming a positive signal are counted by software and the StDev of the total value is divided by a fixed StDev value: 2.5 to calculate the final PC value. These values were analyzed as comparison to the ELISA values for the final analysis of the results.

4.2.7 Custom PA Screening

Custom PAs were cut into individual stripes with a standard paper cutting gillotine and prepared for screenings according to the manufacturer's recommendations. WesternDot 625 Western Blotting kit (Invitrogen) was utilized throughout the entire screening procedure. Custom PAs were agitated in blocking buffer for 1 hour at RT. Each of 50 SCLC and 50 healthy control serum was diluted 1:500 in wash buffer. Each stripe was agitated in a serum dilution for 16 hours at RT. Custom PAs were washed 3 times for 5 minutes in wash buffer with vigorous shaking. Goat anti-human (H+L Specific) ZyMAX Grade (Invitrogen) 2°Antibody is diluted 1:2000 in wash buffer and each stripe was agitated in this solution for 1 hour at RT. PAs were washed 3 times for 5 minutes in wash buffer with vigorous shaking. Q-Dot Conjugate is diluted 1:1000 in blocking buffer and each stripe was agitated in this solution for 1 hour at RT. Custom PAs were washed 3 times for 5 minutes in wash buffer with vigorous shaking. Custom PAs were washed in distilled water for an additional 5

minutes and left for drying on a Whatman paper at RT. Stripes were visualized with StarLight module of Vilber Lourmat Gel Documentation System by utilizing the blue light and red filter of the apparatus. The optical zoom of the camera was adjusted to be maximum and all pictures were taken at this setting.

4.2.8 Custom PA screening results' evaluations

Custom PA screening results were evaluated by MSS and SAS respectively. MSS provided selection of all signals and their sorting from strongest to weakest. SAS provided X values for the signals and based on the cut-off value of 5, positive signals were represented in an Excel sheet showing clone names on one axis and the patients on the other. This Excel sheet was used further for the optimum antigen panel selection based on the custom PA screening results.

4.3 INTER-ASSAY VALIDATION

4.3.1 Insert analysis of pQE plasmids

We have obtained the ordered clones from Imagenes in agar slants and we have inoculated the clones in agar plates to obtain single colonies. We have selected 2 single colonies for each clone for reproducibility and inoculated them in 3 ml LB + Carbenicillin + Kanamycin. The next day we have performed plasmid miniprep for all clones by utilizing Qiagen plasmid miniprep kit. We have selected BamHI and HindIII as the enzymes for double digest of the minipreps for insert analysis and incubated the reaction at 37°C for 3 hours. We run the digested plasmids on 1.5% agarose gel to analyse the insert lengths.

4.3.2 Clone protein expression analysis by IPTG induction and SDS-PAGE

Obtained clones were also analyzed for protein expression. We have inoculated single colonies from each clone into 3 ml LB + Carbenicillin + Kanamycin. The next day we have inoculated 50 microliters of the overnight culture in 10 ml LB and incubated at 37°C shaker until $OD_{600}=0.6$. At this point we have divided the culture into two; one of them was induced by 1mM IPTG and the other one kept as the control. Both cultures were incubated for 4 hours. Both cultures were centrifuged to pellet the E.coli cells and lysed in SDS-PAGE loading buffer and incubated at 100°C for 5 min. Then the samples were run in an 12% SDS-PAGE.

The gel was stained with Coomassie Brilliant blue and the resulting images were analyzed for protein expression induction.

4.3.3 HisTagged Protein Expression and Purification

Each clone was inoculated in LB + Carbenicillin + Kanamycin and grown for 16 hours at 37°C and this culture was used to obtain 15% Glycerol stock for each clone.

For HisTagged protein expression *E. coli* was inoculated from glycerol stock onto agar plate with streak plate technique and grown for 16 hours at 37°C. A single colony from the agar plate is selected and inoculated in 10 ml LB with 50µg/ml of Carbenicillin and 25µg/ml of Kanamycin and grown for 16 hours at 37°C. The culture is poured into 500 ml of LB having the same concentrations of the antibiotics and grown at 37°C until an OD₆₀₀= 0,6 is obtained. Protein expression from the gene of interest was achieved by 1 mM IPTG induction of the culture and the culture was agitated at 37°C for an additional 4 hours. The culture was put on ice for 5 min at the end of the growth interval. The culture was centrifuged at 3500 rpm for 20 min at +4°C. The supernatant was discarded and the pellet was kept at -80°C. The rest of the procedure was performed according to the Qiagen Qiaexpressionist denaturing protein purification protocol and for flow rate consistency during the procedure a peristaltic pump from Ismatec (Model: Reglo MS-4/6) was used. The pure protein's concentration was measured by Invitrogen Qubit Fluorometer.

4.3.4 ELISA

Thermo Scientific Immulon 4HBX ELISA plates were coated with 200 ng/well pure protein in 100 microliters for 16 hours at +4°C. 5% non-fat dried milk in PBS-T was used as the blocking buffer and wells were blocked with 200 microliters/well for 16 hours at +4°C. BioTek ELx405 ELISA washer was used to wash plates with 300 microliters/well of PBS-T for each of 6 repeats by the default settings. Each serum was diluted 1:400 and 1:1600 and each dilution was pipetted as 100 microliters/well in duplicate wells. On each ELISA plate the plate is coated with a single antigen and 4 wells were screened with each serum; 2 x 1:400 and 2 x 1:1600 dilutions, for a total of 24 sera. The final ELISA value for each serum was obtained by averaging all the 4 OD values. Serum incubation was performed for 2 hours at 37°C. Plates were washed as indicated above. Goat anti-human IgG Alkaline Phosphatase conjugated 2° Antibody (Jackson ImmunoResearch) was diluted 1:5000 in blocking buffer and pipetted as 100 microliters/well. 2° Antibody solution was incubated for 2 hours at 37°C.

Plates were washed as indicated above. Thermo Scientific Pierce Phosphatase Substrate Kit was utilized for signal detection. 1 tablet of PNPP was dissolved in 10 milliliters of 1x Diethanolamine Buffer. 100 microliters/well solution was pipetted into wells and incubated for 1 hour at 37°C. OD405 nm was measured by BioTek μ Quant immediately after 100 microliters of 2N NaOH addition into each well.

4.3.5 ELISA Data Normalization

After each ELISA experiment, the newly generated ELISA data was normalized based on 2 independent parameters before addition to the existing data. So the whole ELISA data is a normalized data which means the inter- and intra-experimental differences are eliminated and the data is comparable in itself.

The first level of normalization was performed to eliminate plate to plate differences so that the data coming from different plates becomes comparable. A background value of 0,5 is defined and all the serum OD values from that plate were sorted from the lowest to the highest. Then the lowest 5 values were averaged and 0,5 is subtracted from that average. The remaining value was added to all the values from that plate. The procedure is repeated for each plate independently and plate based abnormalities were eliminated by this strategy.

The second level of normalization was performed to eliminate serum to serum differences so that the data coming from different sera becomes comparable. Again a background value of 0,5 is defined and all the plate based normalized values from a single serum were sorted from the lowest to the highest. Then the lowest 5 values were averaged and 0,5 is subtracted from that average. The remaining value was added to all the values for that serum. The procedure is repeated for each serum independently and serum based abnormalities were eliminated by this strategy.

4.3.6 Sensitivity and Specificity calculation

Sensitivity is calculated as the ratio of seropositive cancer cases at a certain cut-off to all cancer cases. Likewise specificity is calculated as the ratio of seronegative healthy controls for the same cut-off to all healthy controls.

4.3.7 ELISA Evaluations

In our ELISA set-up in each experiment we were able to screen 24 sera for 5-23 different proteins at once. At first the ELISA data was completed for 24 SCLC and 24 Healthy control sera for 27 proteins that were selected based on the PA and custom PA screening results. Then for each of 27 protein's ELISA values the mean (M) and standard deviation (SD) of the healthy control values were calculated. Based on these values multiple cut-offs were calculated as $M+10SD$, $M+7SD$, $M+5SD$, $M+4SD$ and $M+3SD$. The proteins having a sensitivity value at $M+10SD$ were selected first for further evaluation with a new set of 24 SCLC and 24 Healthy control sera. Additional proteins were selected for the next decreasing cut-off value till the number of selected proteins reach 13. We have also screened the 24 SCLC and 24 Healthy control sera with EBV p18 (an immunogenic coat protein) as a positive control and human DHFR as a negative control protein. The number is 15 since routine ELISA screenings with 15 proteins is ideal for practical reasons. The evaluation procedure was repeated with the combined data generated by the addition of the previous data to the new one and the new protein list was defined. Based on this evaluation strategy some of the top 13 proteins retained and some of them were replaced with other proteins from the 27 proteins after each evaluation. This strategy is repeated until the completion of the whole ELISA data of 448 SCLC and 192 Healthy control sera for the top 8 proteins. This data is named as the complete ELISA data and analyzed to obtain the results.

4.3.8 Western blotting

500 ng/well pure protein was separated by 12% NuPAGE Bis-Tris Precast Gel (Invitrogen) under denaturing conditions by utilizing XCell SureLock Mini-Cell (Invitrogen). Wet-transfer of proteins was performed by the transfer apparatus of the same system onto Immobilon-P PVDF Membrane (Millipore) according to manufacturer's recommendations. Membrane was agitated in 5% non-fat dried milk powder in PBS-T for 1 hour at RT for blocking. Serum was diluted 1:500 in blocking buffer and membrane was agitated with serum dilution for 16 hours at RT. Membrane was washed 3 times for 5 minutes in PBS-T with vigorous shaking. Goat anti-human IgG Horseradish Peroxidase 2° Antibody (Jackson ImmunoResearch) was diluted 1:20000 in blocking buffer and membrane was agitated in 2° Antibody solution for 1 hour at RT. Membrane was washed 3 times for 5 minutes in PBS-T with vigorous shaking. ECL Plus Western Blotting Detection Kit (GE Healthcare Amersham) was utilized according to manufacturer's recommendations for signal generation. Signals

were obtained via Vilber Lourmat Gel Documentation System by 5 minutes of camera exposure.

4.3.9 Optimum cut-off determination for ELISA results by MC Method

The optimum ELISA OD cut-off values for each antigen was determined by utilizing a Microsoft Excel Monte Carlo Simulation Add-In. The simulation calculated panel sensitivity and specificity values based on randomly generated cut-offs by rand() function of Excel at each run of the simulation. The simulation was run for 1.000.000 times and at each run the random cut-offs, individual sensitivity and specificity values for each antigen and the panel sensitivity and specificity values were calculated and recorded on an Excel sheet. Then all the rows in this Excel sheet were sorted from the highest panel specificity to the lowest. Then the rows having a specificity >90% were sorted from the highest panel sensitivity to the lowest. Then the highest sensitivity values for each specificity value in the 100-90% specificity range was investigated and these values were used for ROC plotting.

4.3.10 Optimum antigen panel determination from Custom Array Screening Evaluations by MC Method

The optimum antigen panel based on the custom PA screening results was determined by utilizing a Microsoft Excel Monte Carlo Simulation Add-In. The simulation used the Excel sheet generated after custom PA screening results' evaluation and the simulation calculated panel sensitivity and specificity values by combined evaluation of randomly selected antigens which have a sensitivity value, by rand() function of Excel at each run of the simulation. The simulation was run for 1.000.000 times and at each run the antigens which are included in the panel and the panel sensitivity and specificity values were calculated and recorded on different rows in an Excel sheet. Then all the rows in this Excel sheet were sorted from the highest panel specificity to the lowest. Then the rows having a specificity >90% were sorted from the highest panel sensitivity to the lowest. Based on this sorting strategy the highest sensitivity values were recorded for models using the minimum number of antigens for decreasing panel specificity values for the range of 100-90%.

4.4 PROGNOSTIC ANALYSIS OF AUTOLOGOUS ANTIBODIES

All analysis for prognostic correlation of autologous antibodies were performed by utilizing R and GraphPad Prism version 6.00, (GraphPad Software, San Diego California USA), or the Statistical Package for the Social Sciences, version 19 (SPSS Inc., Chicago, IL).

4.4.1 R packages utilized for diagnostic and prognostic analysis

The R packages utilized for diagnostic and prognostic analysis throughout the thesis are; “*libsvm*” for SVM analysis, “*Caret R*” for feature selection and accuracy analysis, “*compound.Cox*” for univariate and multivariate Cox regression, “*limma*” for volcano plots and “*fdrtool*” for cut-off significance for seropositivity and survival correlation analysis.

4.5 ANTIGEN EXPRESSION AND ANTIBODY CORRELATION

4.5.1 Patient and control population

The study cohort consisted of 59 patients with pathologically confirmed SCLC diagnosed between October 2007 and January 2009. All patients gave informed consent and the study was approved by the ethical board of the Atatürk Chest Diseases and Chest Surgery Education and Research Hospital, Ankara, Turkey. All samples were anonymized before analysis. None of the patients had neurological symptoms or evidence of PND within the follow-up period. The clinical data obtained from patients at the time of diagnosis included age, gender, tumor stage, serum alkaline phosphatase (AP), and lactate dehydrogenase levels (LDH). All patients received chemotherapy with or without concurrent and sequential radiotherapy. Survival data was available for all patients. Follow-up times ranged from 0.2 to 44.7 months with a median of 8.68 months. Sera from 157 age-matched healthy controls were obtained after informed consent at the Capa Chest Diseases and Chest Surgery Education and Research Hospital of İstanbul, Turkey.

4.5.2 Enzyme-linked immunosorbent assay

Serum samples were obtained from all 59 patients, collected at the time of diagnosis, between October 2007 and January 2009 and stored at -70°C . SOX2 and two control proteins (EBV-p18 and DHFR) were expressed and purified using the prokaryotic pQE expression system (Qiagen Inc., Valencia, CA, USA). All constructs contained cDNA corresponding to the full ORF of each gene. Ninety-six well Immulon 4 HBX plates (Thermo Scientific, Lafayette, CO, USA) were coated with $0.2\ \mu\text{g}/\text{ml}$ antigen at 4°C for 16 hours. Plates were blocked with 5 % non-fat milk in PBS. Patient sera were added to plates at two dilutions (1:400, 1:1600) and incubated for 2 h at 37°C and subsequently with goat antihuman IgG - AP conjugate (Jackson ImmunoResearch Laboratories Inc., West Grove, PA, USA) at a dilution of 1:5000. The color reaction was read with an ELISA plate reader (BioTek Instruments, Inc., Winooski, VT, USA). Each experiment was repeated twice. A sample was considered seropositive if the average of OD405 values for the two dilutions for a given sample was above the mean plus 2 standard deviation of that obtained for healthy control serum (corresponding to an OD405 value of 1,799). Six control sera (3.8%) were seropositive. DHFR was used as negative and EBV p18 as positive controls. OD values obtained for anti-p18 antibodies from patients and controls were not statistically different.

4.5.3 Western analysis of SOX2 seroreactivity

One hundred nanograms per well of recombinant SOX2 protein was separated by 12% SDS-PAGE under denaturing conditions and transferred to Immobilon-P PVDF membranes (Milipore, St. Charles, MO, USA) using the BioRad semi-dry transfer system. The membranes were then blocked in 5% non-fat milk and incubated with patient serum diluted at 1:3000 for 16 hours at 4°C , after which they were washed in TBST and incubated with goat anti-human IgG (Fc-specific)-HRP (Sigma, St. Louis, MO, USA) for 2 hours. Immunoreactive protein was visualized using ECL-Plus Western Blotting system (GE Healthcare, Buckinghamshire, UK). Mouse anti-human SOX2 monoclonal antibody (R&D Systems, Minneapolis, MN, USA) was used as a positive control.

4.5.4 Immunohistochemistry

Formalin-fixed paraffin embedded tumor tissues from all patients, obtained at the time of diagnosis, were retrospectively evaluated by IHC. Fifty five of 59 tissues had sufficient tumor tissue for a reliable evaluation. Tissues were sectioned at $4\ \mu\text{m}$, placed on positively

charged slides and stained using a Ventana Benchmark LT automatic immunostainer (Ventana Medical Systems, Tuscon, AZ, USA). A range of dilutions of the primary antibody as well as various incubation times and temperatures were tested for optimization. Monoclonal mouse anti-human SOX2 primary antibody (MAB2018, R&D Systems, Minneapolis, MN, USA) was used to stain sections at a 1:25 dilution for 40 minutes. The iViewT DAB Detection Kit (Ventana Medical Systems, Tuscon, AZ, USA), with standard CC1 pretreatment was used for detection. IHC staining was estimated by microscopy as the frequency (percentage) of stained cells as well as by the intensity of staining (graded from 0 to 3). H-scores were calculated as described previously [14]. All samples were evaluated independently by two pathologists (SS and FD). The NTERA-2 cell line was used as a positive control and for optimization experiments. A normal tissue (sausage) block was used as a negative control and was included in every run.

4.5.5 Statistical analysis

Distribution of anti-SOX2 antibody or SOX2 staining in each of age, sex, stage, AP and LDH categories were examined using frequency tables, and differences were evaluated with two sided chi-square tests. Overall survival was defined as the time from diagnosis to death or date of last follow-up. Data for patients that were alive at the last contact were censored. SOX2 protein or antibody effects on survival were estimated by Kaplan-Meier method and the log-rank test was used to compare survival across groups. All data were dichotomized as indicated in the tables and analyzed as categorical variables. All P values were two-sided. All analyses were performed using GraphPad Prism version 6.00, (GraphPad Software, San Diego California USA), or the Statistical Package for the Social Sciences, version 19 (SPSS Inc., Chicago, IL).

4.6 AUTOLOGOUS ANTIBODIES IN AUTOIMMUNE DISEASE

Pool screening of testis and fetal brain PAs and custom PA screening with the clones selected from FB pools screening was also performed for Behçet's disease besides the cancer screenings. The most valuable clones were selected from the pool and individual serum screenings as explained above. The most valuable clones were analyzed for their panel sensitivity and specificity based on the custom PA screening results as performed for cancer screenings.

4.6.1 Identification of autologous antibodies in NBD

The autologous antibody identification methods that are explained above which are developed in the scope of this study was utilized for identification of autologous antibodies specific to neuro-Behçet's disease (NBD). Imagenes fetal brain PAs were screened with pooled NBD sera. 10 NBD sera were pooled and utilized for pool screening. After identification of the strong signals the clones were ordered from Imagenes and protein expression and HisTagged protein purification were performed as explained above. The pure proteins were screened with ELISA and Western blotting for validation of the pool screening results.

4.6.2 ELISA and Western blot validation of autologous antibody levels

ELISA was performed by coating the plates with pure protein and utilizing the serum as the primary antibody. The protocol utilized in the study was shaped by our optimization experiments. The OD values were shown on scatter dot plots with comparison to Behçet's disease, multiple sclerosis, neuromyelitis optica patients and healthy controls for practical visual evaluation. Western blotting was performed by running the pure protein in SDS-PAGE and immunoblotting with the serum as the primary antibody and a Mtch1 antibody as positive control.

Chapter 5

5 RESULTS

5.1 METHOD DEVELOPMENT

Before we do the PA screening we have performed optimization experiments for multiple parameters. These parameters are serum selection for pooling, serum dilution, pool dilution, 2° Ab dilution and Q-Dot dilution.

5.1.1 Serum selection for pooling

Our Western blot results showed that some SCLC sera have non-specific affinity against PVDF membrane (Figure 9). To eliminate undesirable background signals we decided to exclude these sera from pooling.



Figure 9. Non-specific background staining by some SCLC sera (left) and their corresponding ELISA values (right). Each strip corresponds to a single serum. The variable background staining was observed for serum 8 which could contribute non-specific background signal in PA screening. When evaluated by Western blot and ELISA the serum is seronegative.

For exclusion of the sera which will cause undesirable background signal, we have performed an immunostaining with PVDF membrane. We have applied our Western blot protocol to identify the sera causing background staining. The overreactive sera were clarified and excluded from pooling (Figure 10).

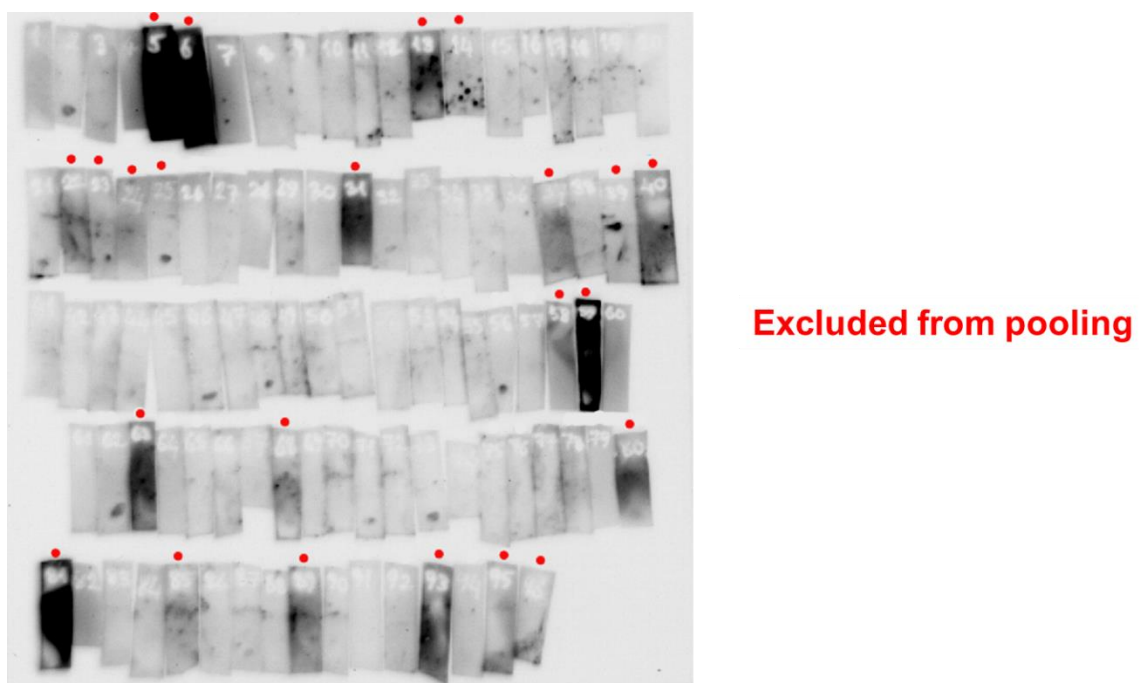


Figure 10. SCLC sera causing undesirable background staining. Each strip corresponds to a single serum. The sera giving a strong non-specific background signal were excluded from pooling. These excluded sera are indicated with a dot.

5.1.2 Serum dilution

To define a serum dilution ratio which will be suitable for detection of high titer antibodies we have performed Western blot against SOX2 with an anti-SOX2 antibody positive serum for the serum dilution ratios: 1:3000, 1:10000 and 1:100000 (Figure 11). The aim was to identify the highest signal/background ratio. 1:10000 was selected as the final concentration of a serum in the pool dilution for ideal signal generation.

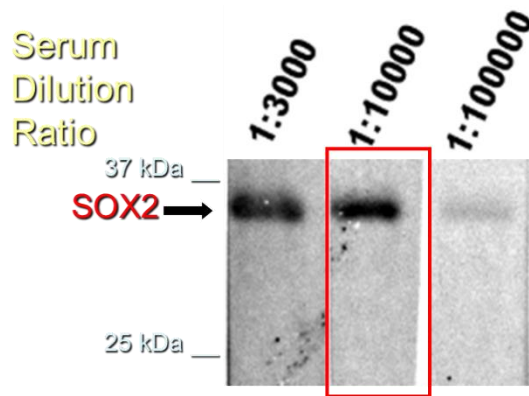


Figure 11. Final serum dilution ratio determination for ideal signal / background ratio. SOX2 was run on SDS-PAGE and immunoblotting was performed with a seropositive serum against SOX2. Among 3 different serum dilutions, the highest signal / background ratio is indicated with a box.

5.1.3 Pool dilution

As the next step for optimization we aimed to define the ideal serum pool dilution ratio for the serum pool. Sera known to be seropositive for p53, HuD and SOX2 were pooled and the pool was diluted in 3 different ratios to obtain 1:5000, 1:10000 and 1:25000 final serum dilution ratios. Both signal/background and cancer/healthy ratios were identified to be maximum for 1:25000 (Figure 12).

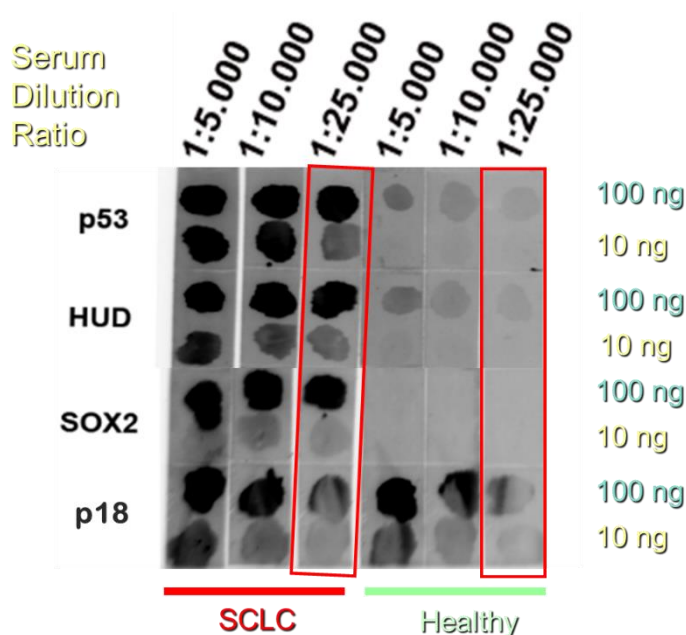


Figure 12. Final serum dilution ratio determination by pool dilution. SCLC and healthy control pools was used as a single serum in DotBlot. Antigens known to elicit autologous antibodies in SCLC; p53, HuD and SOX2 were spotted on PVDF membrane. p18 was also spotted as the positive control. The highest signal / background ratio is indicated with a box.

The lower sensitivity target for the novel antibodies to be identified by PA screening will be 2%. For this sensitivity the number of sera for pooling will be 50, since the rarest autologous antibody will be present in only one patient in this case. So that serum pool obtained from 50 SCLC patients will be diluted 1:200 to obtain 1:25000 final serum dilution ratio.

5.1.4 2° Ab dilution

Two different 2° Ab dilution ratios were tested to identify the ideal one; 1:2000 and 1:500. 1:2000 seems to have less background staining, thus determined as the ideal dilution ratio (Figure 13).

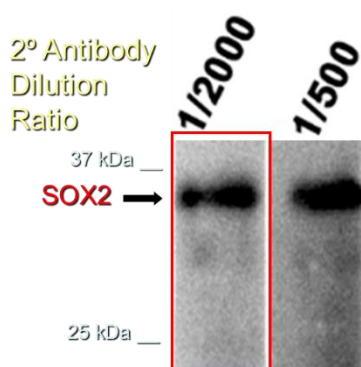


Figure 13. 2° Ab dilution ratio comparison. SOX2 protein and SOX2 seropositive serum also utilized for ideal 2° Ab dilution ratio determination. The highest signal / background ratio is indicated with a box.

5.1.5 Q-Dot dilution

As the final step of protocol optimization 3 different Q-Dot dilution ratios were tested; 1:2000, 1:500 and 1:100. For 1:2000 the signal intensity diminishes whereas for 1:100 there is too much background signal. So 1:500 was determined as the ideal dilution ratio for Q-Dot (Figure 15).

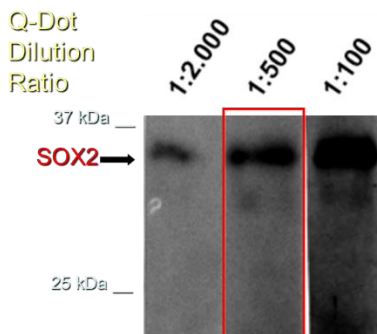


Figure 14. Q-Dot dilution ratio comparison. SOX2 protein and SOX2 seropositive serum also utilized for ideal Q-Dot dilution ratio determination. Among the 3 ratios the highest signal / background ratio is indicated with a box.

5.2 NOVEL AUTOLOGOUS ANTIBODY IDENTIFICATION

5.2.1 Quality Control of PA Screening

We have evaluated the quality control of PA screening by different measures. The first indication that we have done the screening correctly is that we have identified the same antigen from both Testis and FB pool screenings; in this case CIC. We have also identified the

variants of the same antigen on the same pool screening. For example we have identified different variants of SOX2, HuC, EEF1A1, CIC and CHMP5 on FB screening. Whereas we have identified different variants of CCDC9, FN3K, PBXIP1 and YBX1 on the Testis screening. The autologous tumor antigens that were previously identified were reidentified by the screenings like SOX2, HuC and HuD for SCLC, CDR2L and p53 for Ovarian and NY-CO-9 for Colorectal cancer.

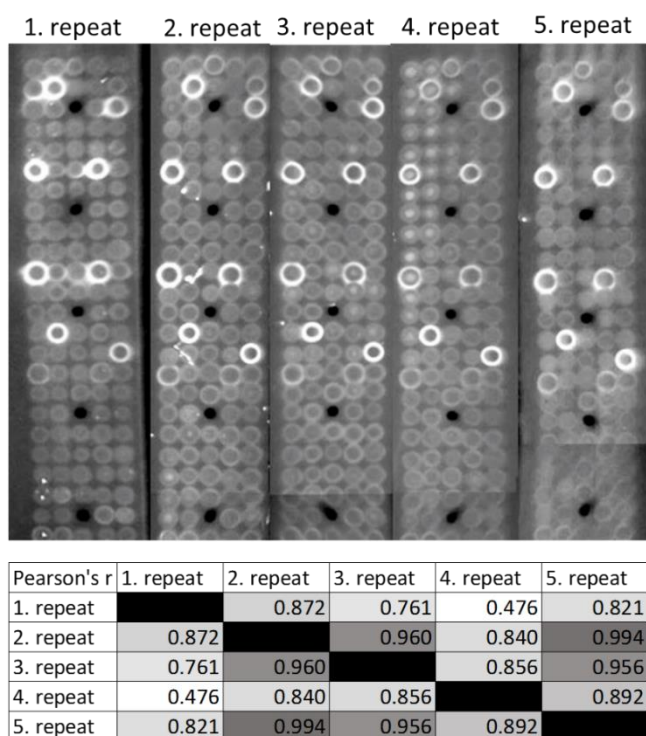


Figure 15. Repeats of custom PA screening by the same serum for reproducibility assessment. The same serum known to give strong, medium and weak signals was utilized to screen a custom PA for 5 different times. Each custom PA screening result is shown (top). Results were converted to numerical values by PC method. Pearson's r correlation values were calculated and shown (bottom).

For quality control of the screening one other parameter was reproducibility of the screenings. To evaluate reproducibility we have screened custom PAs with the same serum for 5 times. Figure 15 shows the reproducibility both visually and with a table including Pearson R values for comparison of repeats. The level of consistency is high based on these values.

5.2.2 Clone selection from pool screenings

After the pool screenings of the Testis PA we have selected the most valuable clones for discriminating disease from healthy controls.

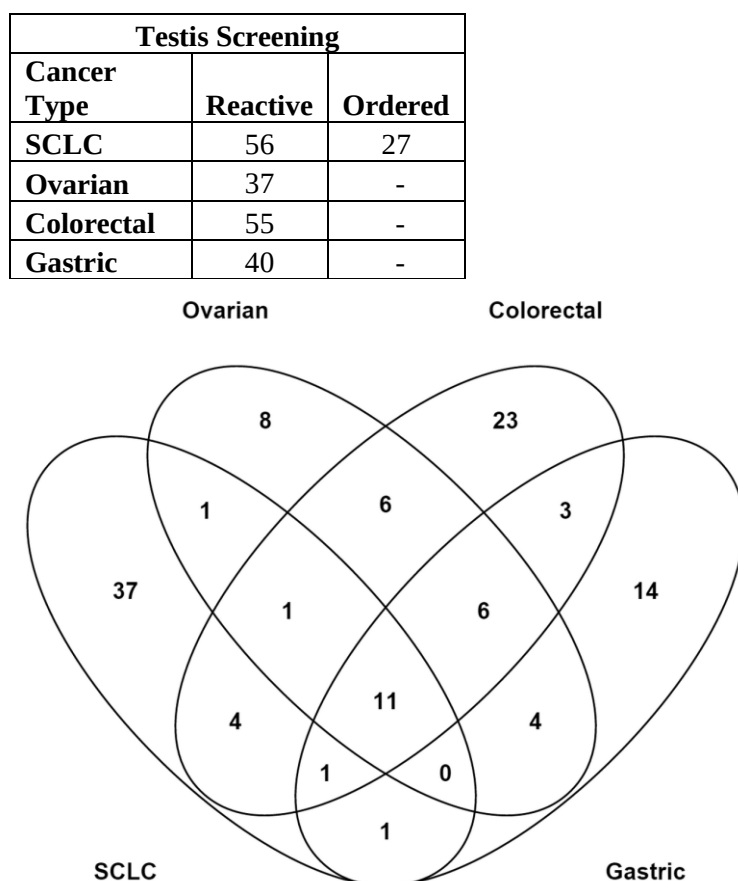


Figure 16. Valuable clones selected from Testis pool screenings and the overlap for different cancers. Reactive clone numbers identified after each pool screening are shown (top). The overlap pattern of the clones from different pool screenings were identified and shown on the venn diagram (bottom). Clones were selected after evaluation with MSS and SAS.

We have identified the number of clones for each cancer screening as shown in Figure 16. We have identified 56 reactive clones by SCLC screening and 37 of them were specific to SCLC. 27 of the 37 clones showed significantly stronger signals for SCLC pool compared to healthy pool and they were selected for further validation with ELISA. On the venn diagram the number of clones reactive for each cancer are shown, some of them were specific to this cancer and some are common to more than one cancer type.

FB Screening		
Cancer Type	Reactive	Ordered
SCLC	22	15
Ovarian	49	-
Colorectal	5	-
Gastric	17	-

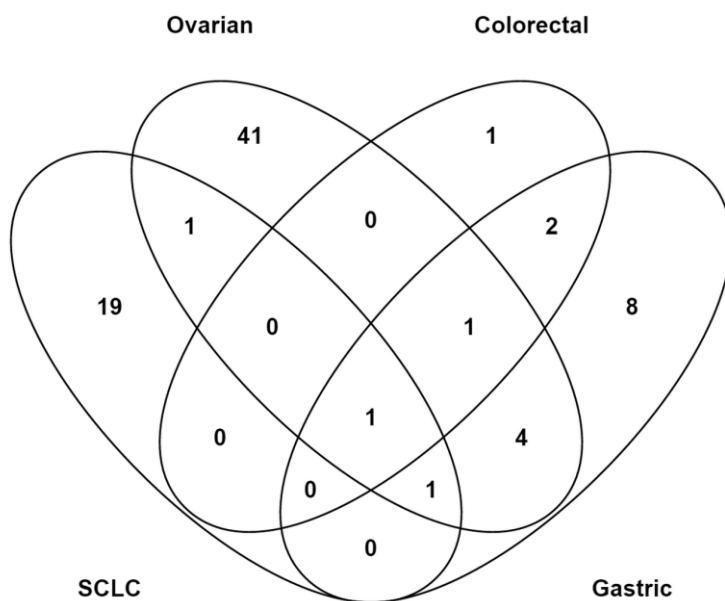


Figure 17. Valuable clones selected from FB pool screenings and the overlap for different cancers. Reactive clone numbers identified after each pool screening are shown (top). The overlap pattern of the clones from different pool screenings were identified and shown on the venn diagram (bottom). Clones were selected after evaluation with MSS and SAS.

The situation is similar for FB screening with pooled sera for the same cancers (Figure 17). Again we have identified 22 clones for SCLC 19 of which are specific for SCLC. 15 of the 19 clones showed significantly stronger signals for SCLC pool compared to healthy pool and they were selected for further validation with ELISA. On the venn diagram again the overlap pattern of these antigens is shown for the pool screenings.

5.2.3 Clones from Testis pool screening

Table 13: List of clones ordered from Testis pool screening

Clones	Clone name	Gene in Clone	Correct Reading Frame
1	SCLC-Testis-16	ALMS1	Yes
2	SCLC-Testis-11	ASRGL1	Yes
3	SCLC-Testis-27	C11orf20	Yes
4	SCLC-Testis-5	CCDC9	Yes
5	SCLC-Testis-12	CCDC9	Yes
6	SCLC-Testis-8	Chr.11 gen.cntg	X
7	SCLC-Testis-21	Chr.19 gen.cntg	X
8	SCLC-Testis-6	CIC	Yes
9	SCLC-Testis-2	FN3K	X
10	SCLC-Testis-3	FN3K	X
11	SCLC-Testis-9	GABBR1	Yes
12	SCLC-Testis-10	NOC4L	Yes
13	SCLC-Testis-14	NUP155	X
14	SCLC-Testis-13	PBXIP1	Yes
15	SCLC-Testis-26	PBXIP1	X
16	SCLC-Testis-24	PTGDS	Yes
17	SCLC-Testis-1	SF3A2	X
18	SCLC-Testis-7	TUSC1	Yes
19	SCLC-Testis-17	UCKL1	X
20	SCLC-Testis-19	WDR90	X
21	SCLC-Testis-4	YBX1	Yes
22	SCLC-Testis-20	YBX1	Yes
23	SCLC-Testis-22	YBX1	Yes
24	SCLC-Testis-23	YBX1	Yes
25	SCLC-Testis-15	YBX2	Yes
26	SCLC-Testis-18	ZFPM1	X
27	SCLC-Testis-25	ZNF160	X

Table 13 shows the clones we have ordered after the Testis PA screening which are 27 in total. As seen on the table the ones indicated with gray shade include variants of the same protein. It is also interesting that some of the clones don't contain a cDNA clone in the correct frame with the expression vector whereas they are identified as recognized by autologous antibodies specific to SCLC. The number of distinct non-redundant proteins from all these clones are 12 in total.

5.2.4 Clones from FB pool screening

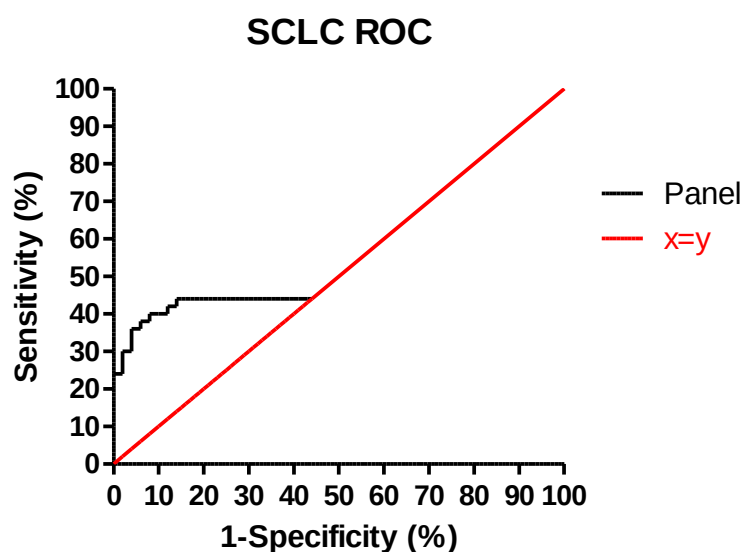
The clones selected from FB PA are utilized to be spotted on custom PAs for screening with individual sera. A number of clones from the pool screenings were selected based on seropositivity strength. For this purpose 40 clones from SCLC, 68 clones from Ovarian, 30 clones from Gastric and 7 clones from Colorectal were selected to be spotted on custom PA. Each custom array was screened with an individual serum and results were evaluated to define sensitivity and specificity for each antigen and the panels of antigens.

Table 14: List of clones ordered from FB pool screening

Clones	Clone name	Gene in Clone	Correct Reading Frame
1	SCLC-FB-1	OTUD5	Yes
2	SCLC-FB-2	HNRPD	Yes
3	SCLC-FB-3	DUSP8	Yes
4	SCLC-FB-4	POLB	Yes
5	SCLC-FB-5	SELO	Yes
6	SCLC-FB-6	PRDM8	Yes
7	SCLC-FB-7	ANKIB1	Yes
8	SCLC-FB-8	BTBD7	Yes
9	SCLC-FB-9	RMND5B	Yes
10	SCLC-FB-10	RNASEN	Yes
11	SCLC-FB-11	hETEF1 α 2	Yes
12	SCLC-FB-12	RPS27A	Yes
13	SCLC-FB-13	HDAC5	Yes
14	SCLC-FB-14	QTRT1	Yes
15	SCLC-FB-15	HIP1R	Yes

The clones identified and ordered from FB PA screening are shown in Table 14. All clones are confirmed to be in frame with the bacterial expression vector. The list is finalized after eliminating redundancy.

5.2.5 Sensitivity and Specificity of custom PA results

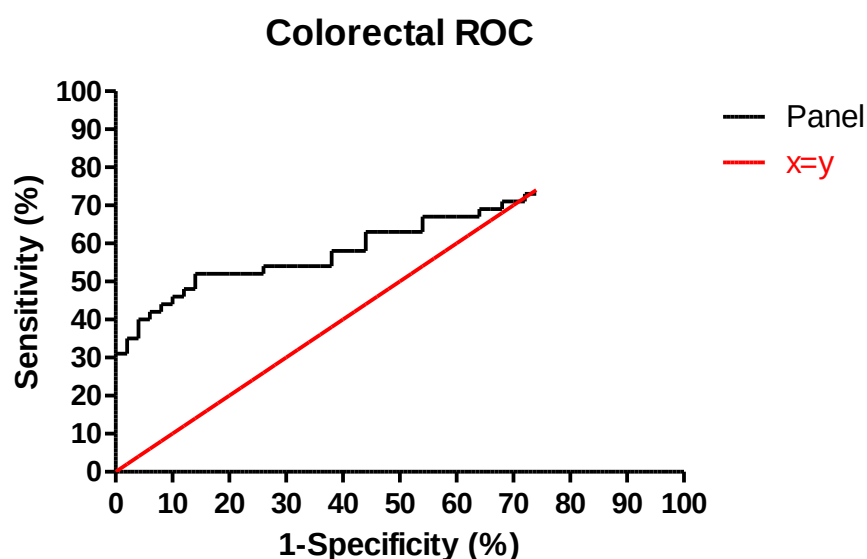


Sensitivity(%)	Specificity(%)	Unique Antigens
24	100	5
30	98	6
36	96	7
38	94	7
40	92	8
42	88	8

Figure 18. ROC for SCLC by PA values evaluated by MC method. SAS values of custom PA results were utilized for MC analysis. MC outputs were analyzed for selection of the best panels giving the highest sensitivity value at each specificity value with the least number of clones. These values were utilized for plotting the ROC (top). Representative sensitivity and specificity values forming the ROC and the corresponding clone numbers in each panel was shown (bottom).

We have evaluated the custom PA screening results and converted the positivities to numeric values. These numeric values were analyzed by MC algorithm and panels were generated randomly by excel. Later on these panels were evaluated according to their sensitivity and specificity values and the panels with the highest sensitivity and minimum number of antigens at each specificity value were identified. The ROC curve was generated by integration of all these panels. For example the best panel had 24% sensitivity with 5

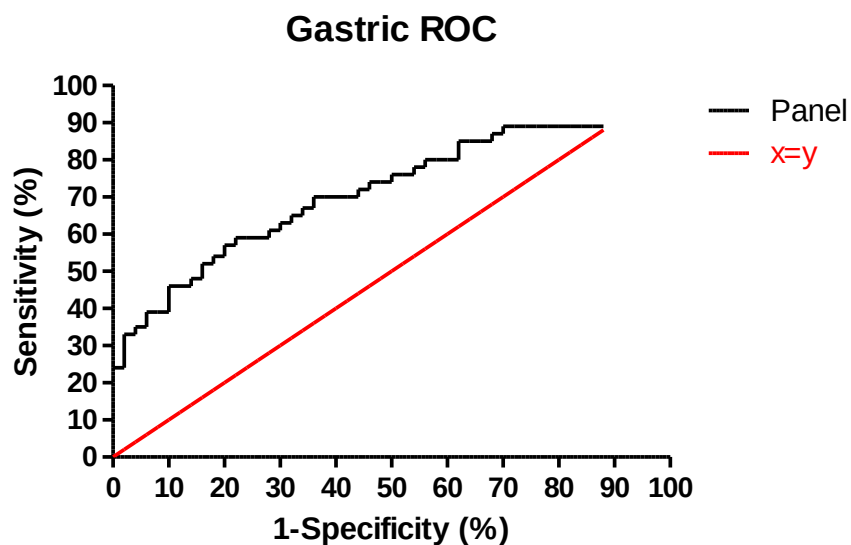
antigens at 100% specificity. The sensitivity increases to 42% with a 8 antigen panel when the specificity decreases to 88% (Figure 18).



Sensitivity(%)	Specificity(%)	Unique Antigens
31	100	2
35	98	6
40	96	8
42	94	6
44	92	8
46	90	6

Figure 19. ROC for Colorectal cancer by PA values evaluated by MC method. SAS values of custom PA results were utilized for MC analysis. MC outputs were analyzed for selection of the best panels giving the highest sensitivity value at each specificity value with the least number of clones. These values were utilized for plotting the ROC (top). Representative sensitivity and specificity values forming the ROC and the corresponding clone numbers in each panel was shown (bottom).

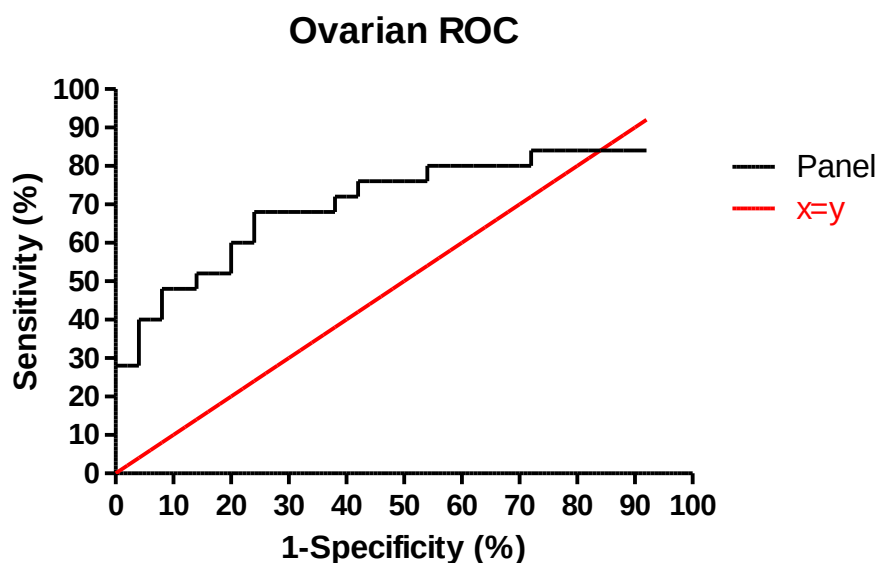
The situation for colorectal cancer is as following: 31% sensitivity with 2 antigens at 100% specificity and the sensitivity value increases to 46% with 6 antigens at 90% specificity (Figure 19).



Sensitivity(%)	Specificity(%)	Unique Antigens
24	100	9
33	98	8
35	96	10
39	94	6
39	92	6
46	90	10

Figure 20. ROC for Gastric cancer by PA values evaluated by MC method. SAS values of custom PA results were utilized for MC analysis. MC outputs were analyzed for selection of the best panels giving the highest sensitivity value at each specificity value with the least number of clones. These values were utilized for plotting the ROC (top). Representative sensitivity and specificity values forming the ROC and the corresponding clone numbers in each panel was shown (bottom).

The same analysis was also performed for custom array screening with Gastric cancer. In this case the best panel has a sensitivity value of 24% with 9 antigens at 100% specificity. Whereas the sensitivity value increases to 46% with 10 antigens at 90% specificity (Figure 20).



Sensitivity(%)	Specificity(%)	Unique Antigens
28	100	6
28	98	8
40	96	10
40	94	10
48	92	10
48	90	10

Figure 21. ROC for Ovarian cancer by PA values evaluated by MC method. SAS values of custom PA results were utilized for MC analysis. MC outputs were analyzed for selection of the best panels giving the highest sensitivity value at each specificity value with the least number of clones. These values were utilized for plotting the ROC (top). Representative sensitivity and specificity values forming the ROC and the corresponding clone numbers in each panel was shown (bottom).

For ovarian cancer the best panel had 28% sensitivity with 6 antigens at 100% specificity and at 90% specificity the best panel has 48% sensitivity with 10 antigens (Figure 21).

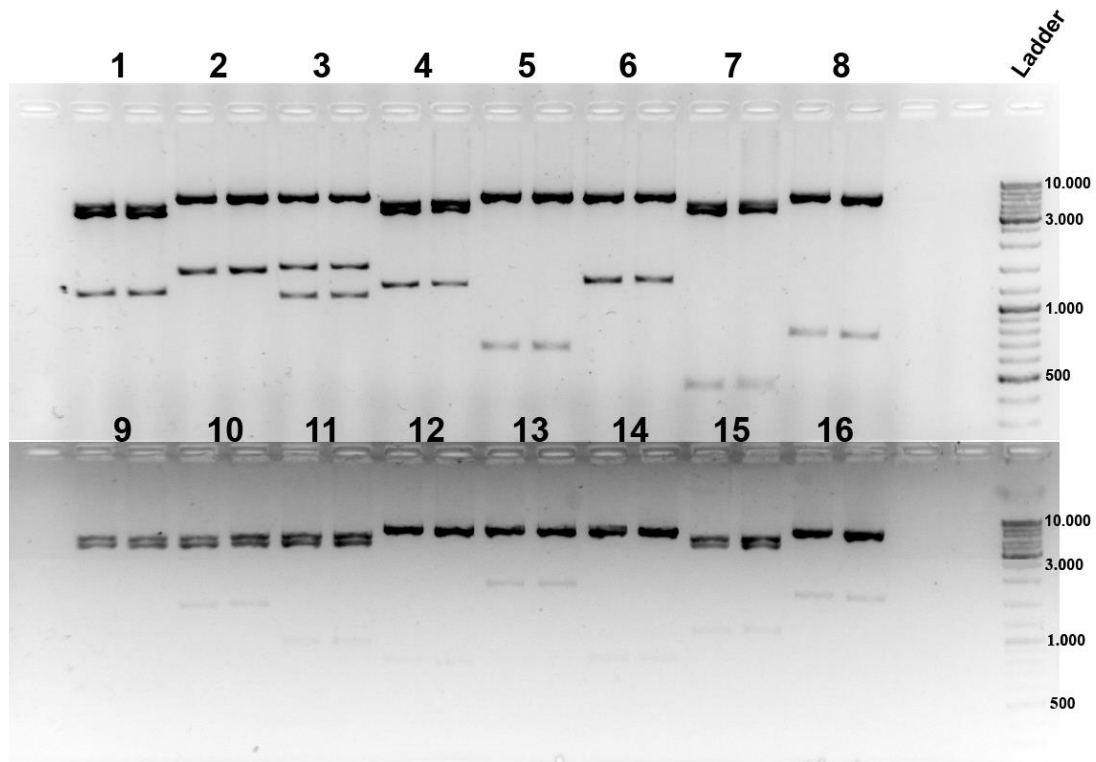
Table 15: Final list of clones screened with ELISA

No	Clone	Gene info	Protein Identity of clone
1	SCLC-Testis-4	YBX1 (NM_004559.3)	276 aa (1-276) of the 324 aa
2	SCLC-Testis-5	CCDC9 (NM_015603.2)	72 aa (452-523) of the 531 aa
3	SCLC-Testis-6	CIC (NM_015125.3)	164 aa (1445-1608) of the 1608 aa
4	SCLC-Testis-7	TUSC1 (NM_001004125.2)	94 aa (54-147) of the 212 aa
5	SCLC-Testis-9	GABBR1(NM_021904.2)	126 aa (249-374) of the 844 aa
6	SCLC-Testis-11	ASRGL1 (NM_025080.3)	196 aa (1-196) of the 308 aa
7	SCLC-Testis-13	PBXIP1 (NM_020524.2)	493 aa (1-493) of the 731 aa
8	SCLC-Testis-15	YBX2 (NM_015982.3)	259 aa (1-259) of the 364 aa
9	SCLC-Testis-16	ALMS1 (NM_015120.4)	449 aa (1316-1764) of the 4169 aa
10	SCLC-Testis-27	C11orf20 (NM_001039496.1)	197 aa (4-200) of the 200 aa
11	SCLC-FB-1	OTUD5 (NM_001136157.1)	241 aa (326-566) of the 566 aa
12	SCLC-FB-2	HNRPDL (NM_001207000.1)	274 aa (90-363) of the 363 aa
13	SCLC-FB-4	POLB (NM_002690.2)	99 aa (237-335) of the 335 aa
14	SCLC-FB-6	PRDM8 (NM_001099403.1)	235 aa (194-428) of the 689 aa
15	SCLC-FB-7	ANKIB1 (NM_019004.1)	900 aa (190-1089) of the 1089 aa
16	SCLC-FB-8	BTBD7 (NM_001002860.2)	208 aa (925-1132) of the 1132 aa
17	SCLC-FB-10	DROSHA (NM_001100412.1)	149 aa (1189-1337) of the 1337 aa
18	SCLC-FB-12	RPS27A (NM_002954.5)	156 aa (1-156) of the 156 aa
19	SCLC-FB-14	QTRT1 (NM_031209.2)	35 aa (47-81) of the 403 aa
20	SCLC-1	SOX2 (NM_003106.3)	307 aa (6-312) of the 317 aa
21	SCLC-2	NY-ESO-1 (NM_139250.1)	112 aa (10-121) of the 180 aa
22	SCLC-3	p53 (NM_000546.5)	393 aa (1-393) of the 393 aa
23	SCLC-4	HuD (NM_021952.3)	370 aa (11-380) of the 380 aa

5.3 INTER-ASSAY VALIDATION

5.3.1 Insert Analysis of pQE plasmids

After we obtained the clones we have performed miniprep plasmid purification and restriction enzyme digestion to each clone in duplicate for insert length analysis (Figure 22).



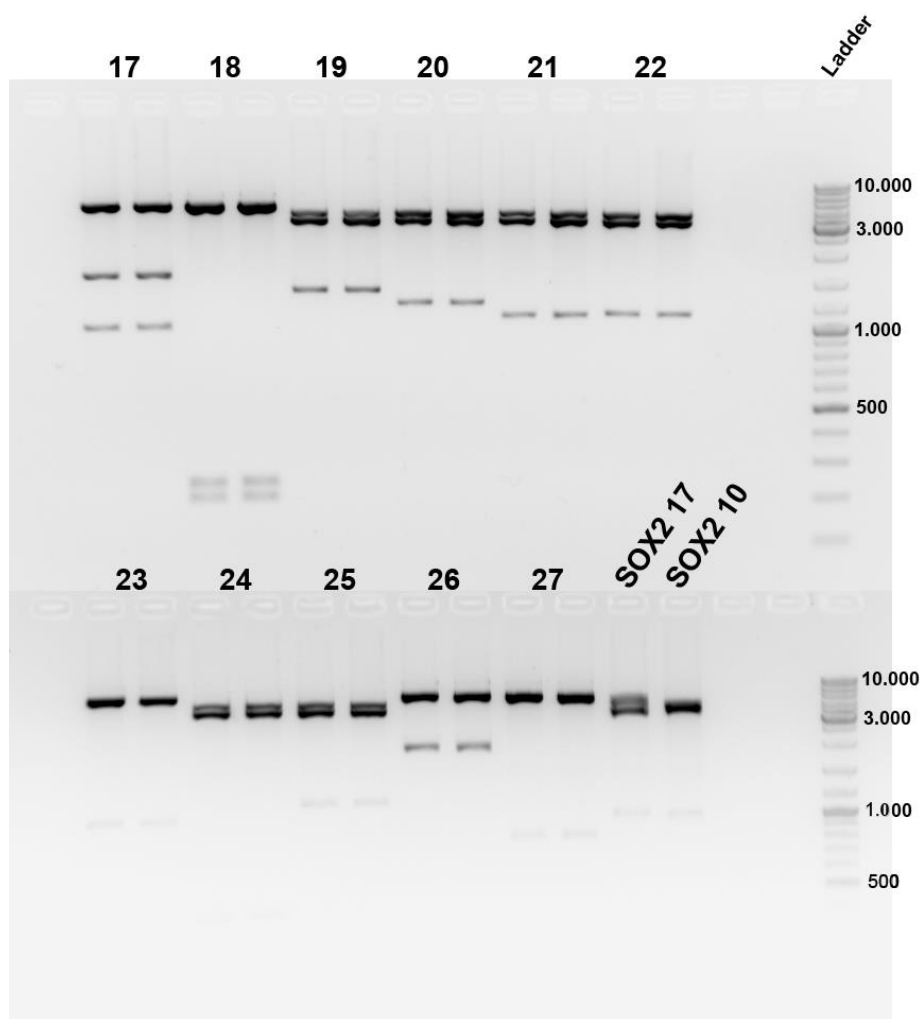


Figure 22. Insert length analysis by restriction enzyme digestion for the clones ordered from Testis pool screening. Each number represents the number in the clone names (for example 1 means SCLC-Testis-1). 2 different colonies from the agar plate was analyzed for each clone for confidence. Plasmid miniprep was performed to 3 ml of overnight culture of each clone. 1 μ g of plasmid DNA was incubated for 2 hours with BamHI and HindIII and run on an agarose gel. Both vector bands and insert bands are visible on the gel. Sections of the images were adjusted for contrast for easier visualization of insert bands.

We have observed both pQE30 (3494 bp) and pQE80 (4806 bp) plasmids as their lengths differ. The inserts varied dramatically (between 200 and 2000 bp) as observed which is an indication of enriched protein diversity represented by the PA.

5.3.2 IPTG induction control of the selected clones

Clones were also analyzed for their HisTagged protein expression capacity. When the E.coli lysates were analyzed by SDS-PAGE and Coomassie blue staining most of the clones were observed not to be induced by IPTG. There was no observable band indicating induction. Only 7 of 27 clones were capable of protein expression due to IPTG induction (Figure 23).

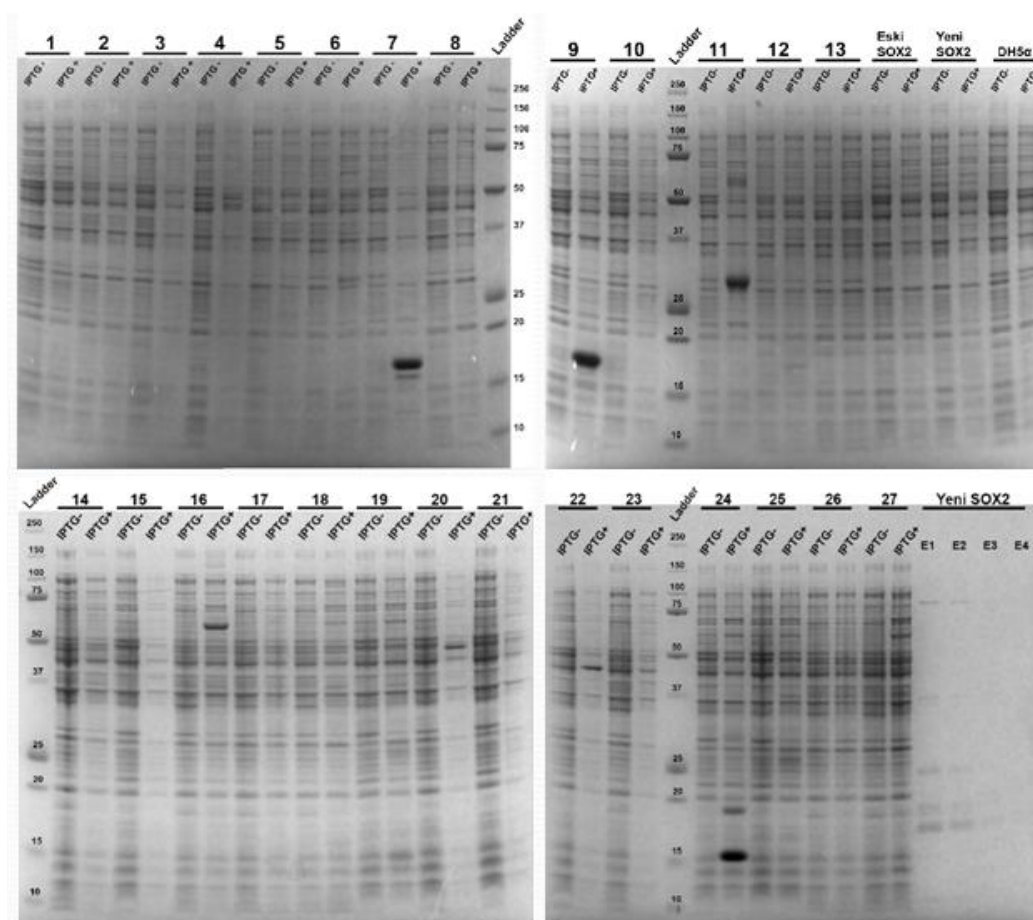


Figure 23. IPTG induction analysis of clones ordered from Testis pool screening. HisTagged protein expression was induced by 1mM final IPTG concentration when the E.coli $OD_{600}=0.6$. Clones were incubated in 37°C for an additional 4 hours. IPTG – lanes correspond to the uninduced culture of the same clone. Clones were pelleted at 3500 rpm for 20 min and the E.coli pellets were lysed by SDS-PAGE loading buffer. Clone numbers are indicated above each lane.

However later on we have realized that the lack of induction was due to lower detection limit of Coomassie blue staining (results are not shown). When grown in 500 ml LB

almost all clones having a true ORF for the cDNA they include were capable of expressing a HisTagged protein in the expected molecular weight.

5.3.3 Iterative ELISA Approach

After elimination of redundancy and out-of-frame clones finally a list of 23 clones both from Testis and FB pool screenings were selected for ELISA (Table 15). Induced protein expression and HisTagged affinity purification was performed for these clones to obtain pure proteins for ELISA.

ELISA Strategy

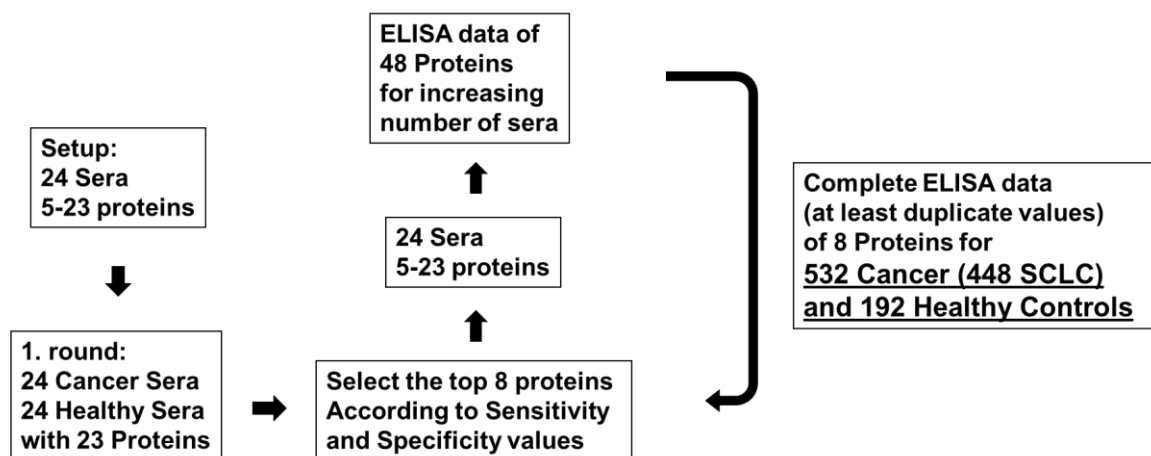


Figure 18. Our iterative ELISA strategy for identification of best antigens by performing the minimum number of screening. The strategy starts with screening the appropriate number of proteins with 24 cancer and 24 healthy sera. The most valuable 8 proteins were selected and the screening with 24 sera was repeated. The accumulated results were evaluated and the most valuable 8 proteins were reselected. This iterative approach was applied till all the sera were screened for the lastly

We have applied an iterative ELISA strategy to the 23 proteins. While the number of sera increased for these 23 proteins we excluded the proteins without any discriminative power. So the data was completed for 8 proteins by screening 448 SCLC and 192 healthy sera in total (Figure 24). This approach was adapted as the most economical method for evaluation of all sera for the most valuable proteins.

5.3.4 Stepwise clone selection by ELISA evaluations

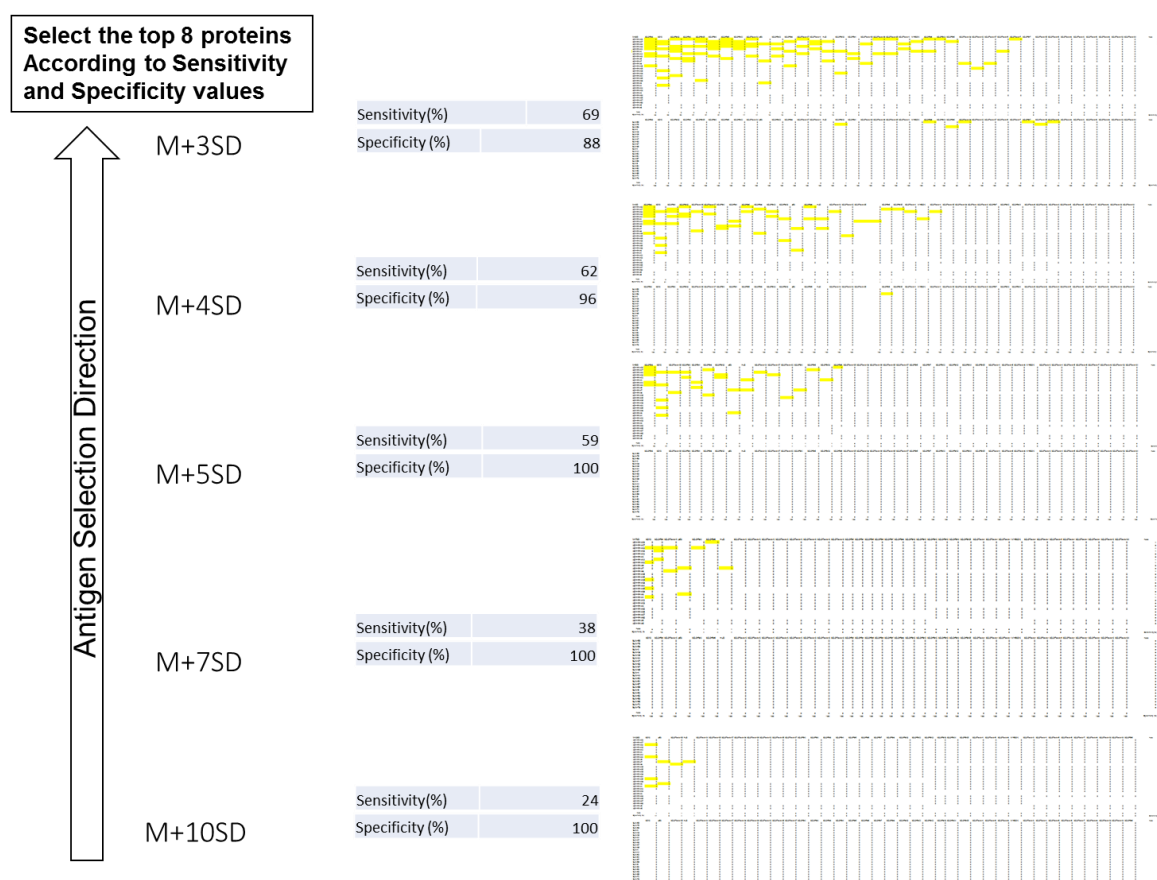


Figure 19. Stepwise antigen selection strategy by utilizing standart deviation cut-offs. Representative sensitivity and specificity tables are shown for each cut-off. The seropositivities were indicated by yellow shade. The vertical arrow indicates the antigen selection direction till the number of the selected antigens is 8.

After the normalization ELISA data was evaluated according to 5 different SD cut-offs. These cut-offs were defined by calculating the average and SD of the healthy controls for each protein. The cut-offs were M+10SD, 7SD, 5SD, 4SD and 3SD. The ELISA values above cut-off were marked with yellow for each cut-off at the corresponding tables both for SCLC and healthy control sera (Figure 25). The antigens which have the highest sensitivity values at M+10SD were selected first and then M+7SD and so on until a total of 8 proteins were obtained after each cycle of ELISA till the completion of the data for all sera.

5.3.5 Normalization for ELISA data

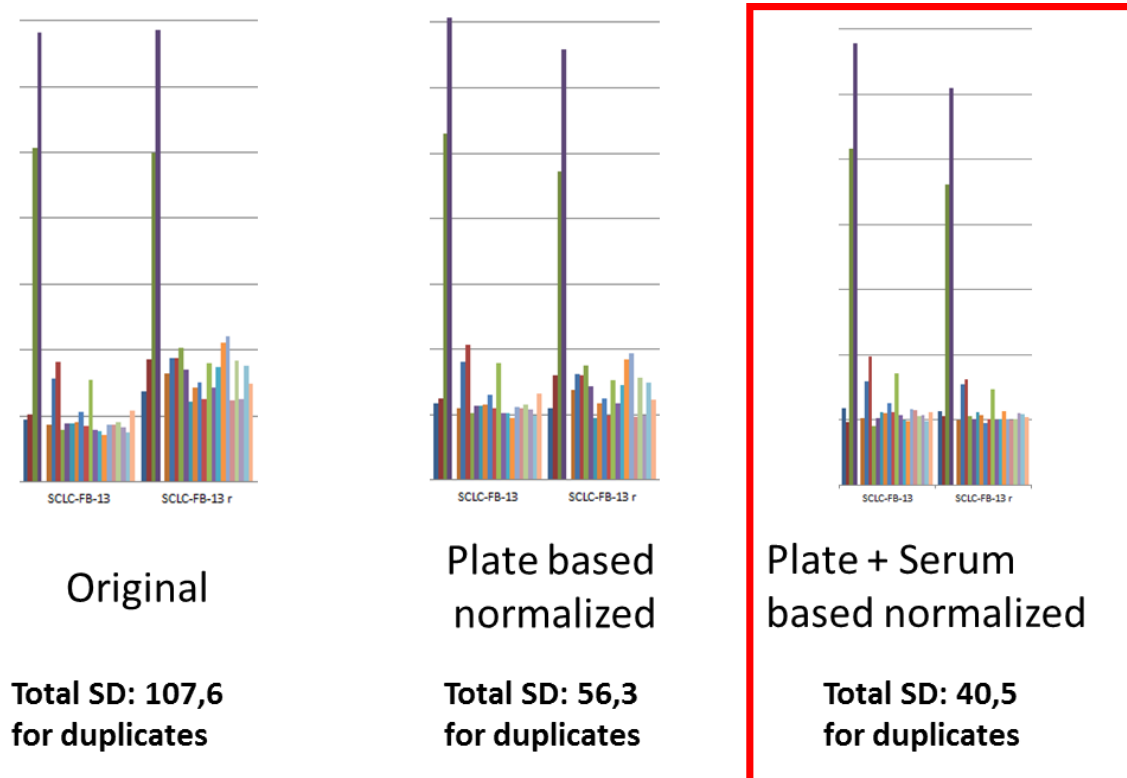


Figure 20. ELISA data normalization for plate and serum. The bar plots show the ELISA values of replicates representatively for SCLC-FB-13. Each bar represents a distinct serum and the order of sera is the same for replicates for easy comparison. Each horizontal bar indicates 0.5 OD increase. Original (left) shows the unnormalized values, plate based (middle) and plate + serum based normalized values (right) are shown.

After completing the ELISA data we normalized the values according to 2 different parameters. The first level of normalization was plate based normalization to eliminate bias from plate to plate. And the second level of normalization was serum based normalization to eliminate bias from serum to serum. To evaluate the effect of these 2 normalizations on the ELISA data we have analyzed a replicate ELISA experiment. In that experiment a single protein was screened for 24 different sera 2 times. The standard deviations for each replicate pair were calculated and the sum of all SD values were calculated. These values were 107.6, 56.3 and 40.5 for the non-normalized, plate based normalized and plate + serum based normalized values respectively (Figure 26). These values clearly indicate the benefits of the 2 levels of normalizations on the ELISA values. Also it is obviously seen that after normalization some weak positive values become more reproducible. So that these 2 normalizations increase the lower sensitivity limit of the assay.

5.3.6 Coefficient of Variation (CV) Analysis

Table 16: CV for ELISA data as quality control (A: Medium Signals)

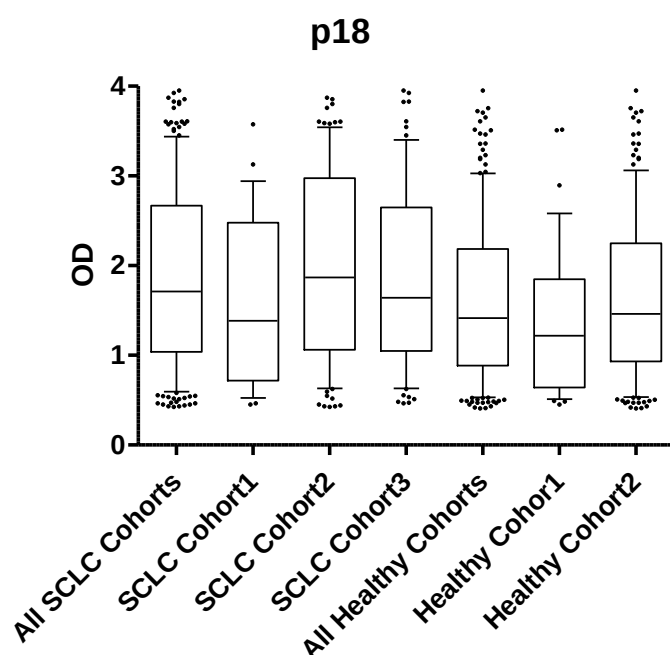
CV (%)	SCLC-1	SCLC-2	SCLC-3	SCLC	Healthy-1	Healthy-2	Healthy-3	Healthy
SOX2	12	13	18	14	13	12	14	13
NY-ESO-1	11	9	16	11	13	18	17	17
p53	14	10	15	12	10	11	11	11
POLB	13	8	13	10	8	7	8	8
ALMS1	10	8	11	9	11	7	9	8
TUSC1	11	7	8	8	8	7	9	8
C11orf20	20	16	22	18	12	19	16	17
HNRPD	15	13	15	14	8	9	12	10
# of Sera	96	244	108	448	35	100	57	192

Table 17: CV for ELISA data as quality control (B: High Signals)

CV (%)	SCLC-1	SCLC-2	SCLC-3	SCLC	Healthy-1	Healthy-2	Healthy-3	Healthy
SOX2	6	4	8	5	42	36	22	30
NY-ESO-1	18	22	13	19	53	55	18	36
p53	4	11	6	8	51	22	23	23
POLB	48	15	52	40	11	21	17	18
ALMS1	28	25	34	30	11	24	24	22
TUSC1	21	28	40	30	23	13	36	24
C11orf20	15	15	25	18	3	26	15	19
HNRPD	16	41	26	30	10	15	26	19
# of Sera	96	244	108	448	35	100	57	192

ELISA data was also evaluated for reproducibility by CV calculation. CV was calculated for replicate values both for medium signals below $M+3SD$ and above that cut-off. These CV values are compared to the inter-experimental and intra-experimental CV calculated for the EarlyCDT-Lung data. Overall comparison shows that our results have lower CV values compared to EarlyCDT-Lung both for medium signals (Table 16) and high signals (Table 17) which indicates that our ELISA data is statistically significantly reliable [26].

5.3.7 Quality control for serum integrity



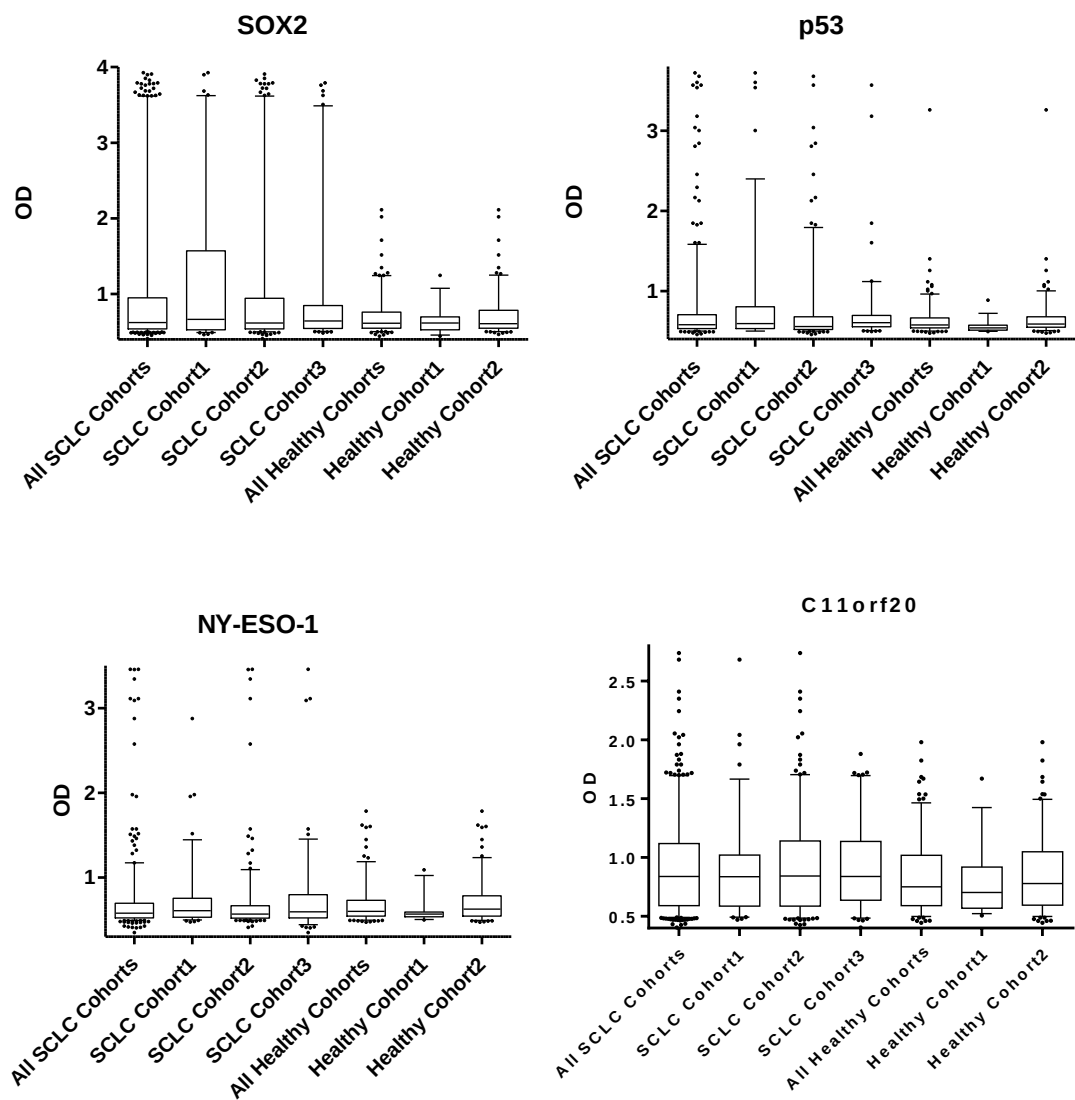
Groups to be compared	p value
SCLC Cohort1 vs All Healthy Cohorts	0.7739
SCLC Cohort1 vs Healthy Cohort1	0.2616
SCLC Cohort1 vs All SCLC Cohorts	0.2769
SCLC Cohort1 vs SCLC Cohort3	0.4232

Figure 21. Quality control for the serum antibody integrity. Normalized ELISA values for p18 were compared for all cohorts. The scatter dot plot shows the OD values for each cohort. Boxes indicate the 50 percentile while whiskers indicate 10-90 percentiles, horizontal lines indicate median values (top). Groups were compared by t-test and the p-values are shown (bottom).

As another quality control measure we also analyzed the integrity of the antibodies in serum. To do that we have screened all sera against an EBV viral coat antigen p18 and the population has around 70% seropositivity level to this antigen. Normally more than 98% of the population is infected by EBV at the age of 11 and elicit antibodies to several antigens of this virus. When we compare our cohorts for the seropositivity against this antigen we see no statistically significant difference between our screening cohorts (Figure 27). This data

suggests that there is no bias related to the integrity of antibodies in sera of any of these cohorts.

5.3.8 ELISA Data



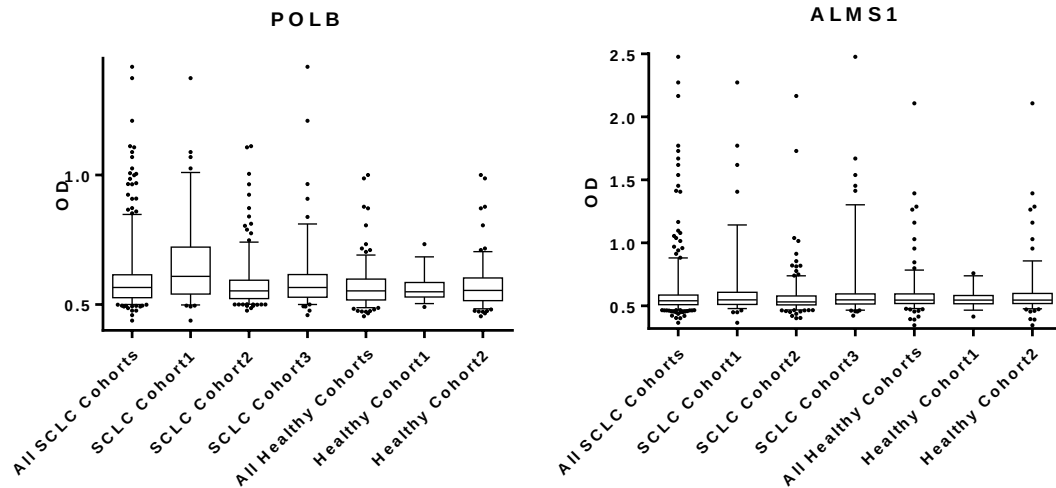


Figure 22. Scatter dot plot of ELISA values of the most valuable 6 antigens. Normalized ELISA values for the most valuable 6 antigens were shown. The scatter dot plot shows the OD values for each cohort. Boxes indicate the 50 percentile while whiskers indicate 10-90 percentiles, horizontal lines indicate median values.

The scatter dot plots show the ELISA values after normalization of the most valuable 6 of the 8 antigens for all cohorts (Figure 28). Obviously SCLC sera have higher seropositivity values compared to the healthy controls for all these antigens. These ELISA values were analyzed by performing different methods for cut-off determination and panel formation.

5.3.9 Manual cut-off determination

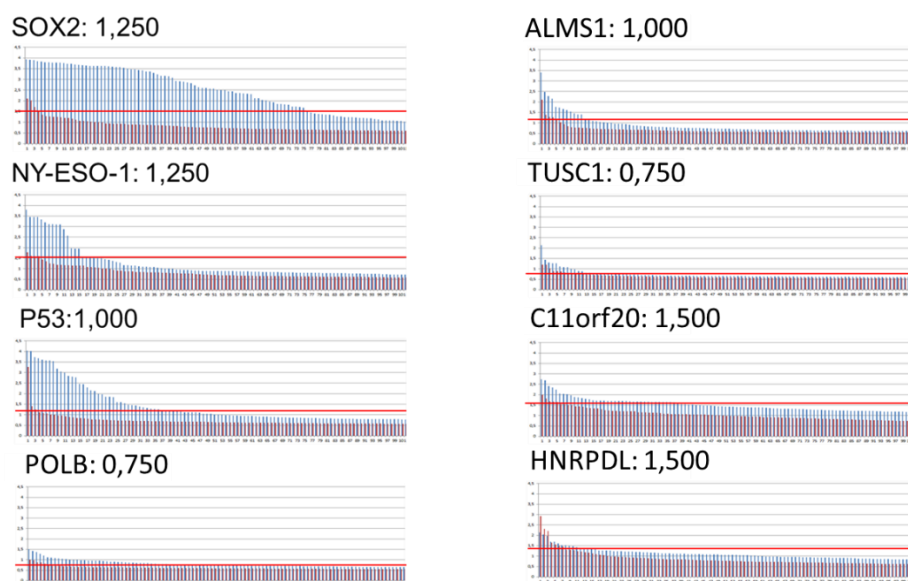


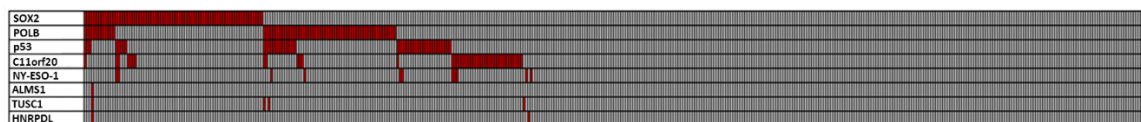
Figure 23. Manual cut-off determination for seropositivity for the ELISA results.

Both cancer and healthy ELISA values for each antigen was sorted from highest to lowest for each group separately. Red bars show ELISA values for each protein for the cancer cohorts while blue bars for the healthy cohorts. A manual horizontal line was plotted for each antigen indicating the manual cut-off. The cut-off were utilized for sensitivity and specificity calculations.

For evaluation of these results one of the strategies for cut-off determination was specifying a manual cut-off based on analysis of the ELISA values in a sorted trend. In these graphs all the 8 antigens are shown and the values are sorted from highest to lowest. The blue bars show the SCLC values whereas the red bars show the healthy control values. The horizontal red lines show the manual cut-off values for each antigen. The cut offs were defined in order to maximize both sensitivity and specificity (Figure 29).

5.3.10 Cut-off determination by Monte Carlo

Panel Sensitivity (%): 42



Panel Specificity (%): 90

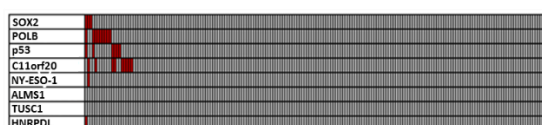
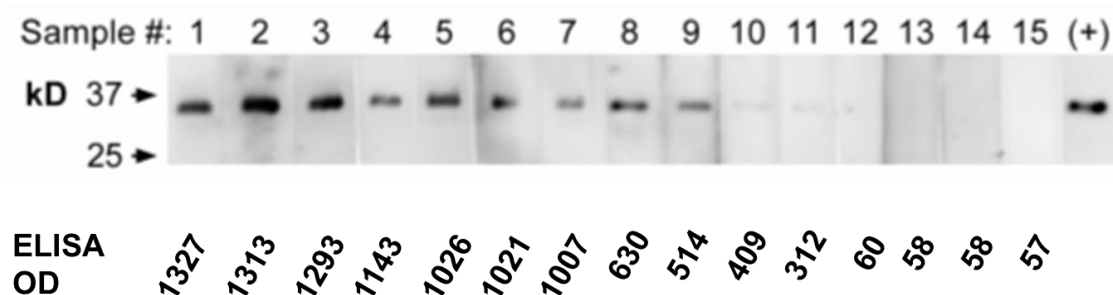


Figure 24. Cut-off determination utilizing the MC algorithm. MC was utilized to define the best cut-offs for all the 8 antigens. The graphs show the seropositivity for all antigens simultaneously for the MC cut-offs. The x axis indicates the patients. The panel sensitivity and specificity values are indicated. The graphs include all cohorts both for SCLC and healthy controls.

One other strategy that we have utilized for cut-off determination for the ELISA values is the Monte Carlo method. In this method the algorithm generates random cut-offs for each antigen independently for 1.000.000 times. Each time sensitivity and specificity of each antigen and the panel is calculated and these values are stored by excel. Afterwards the values are sorted in excel and the highest sensitivity values for each specificity value are identified. Figure 30 shows the seropositives in red for each antigen at the best cut-off combination which gives 42% sensitivity at 90% specificity for the panel. There is a significant amount of overlap of seropositivity among antigens whereas there is a group of patients who are seropositive for only one antigen both for the SCLC patients and the healthy controls which is expected.

5.3.11 ELISA and Western blot comparison



Cut-off ELISA OD for SOX2 = 311

Figure 25. Confirmation of SOX2 ELISA with Western blot for a representative group of sera. Each strip shows the seropositivity of a SCLC patient against SOX2. Corresponding ELISA values are shown. + control is an RGSHis antibody. 311 is the ELISA seropositivity cut-off for SOX2 antibodies.

We have also performed validation control experiments to evaluate the reliability of our ELISA data. To do that we have tried to confirm the ELISA values with Western blotting for a representative group of SCLC patients. There is a high level of correlation between the ELISA and Western blot results. We have sorted the ELISA results from highest to lowest and compared them with their corresponding Western blot results (Figure 31). The rightmost lane represents the positive control for Western blot. The cut-off for seropositivity of ELISA is set to be M+3SD in this case which is 311 as an OD value.

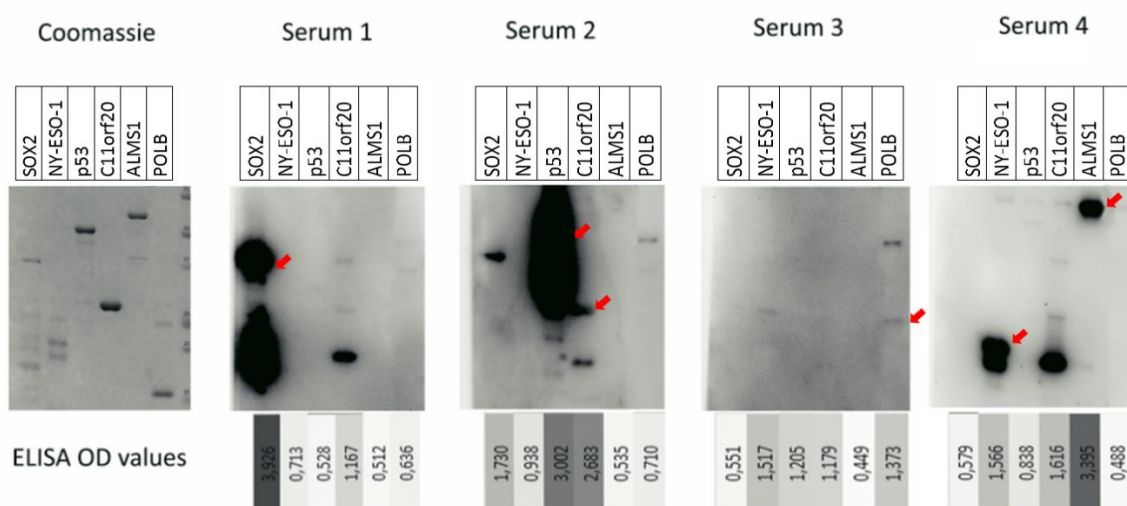
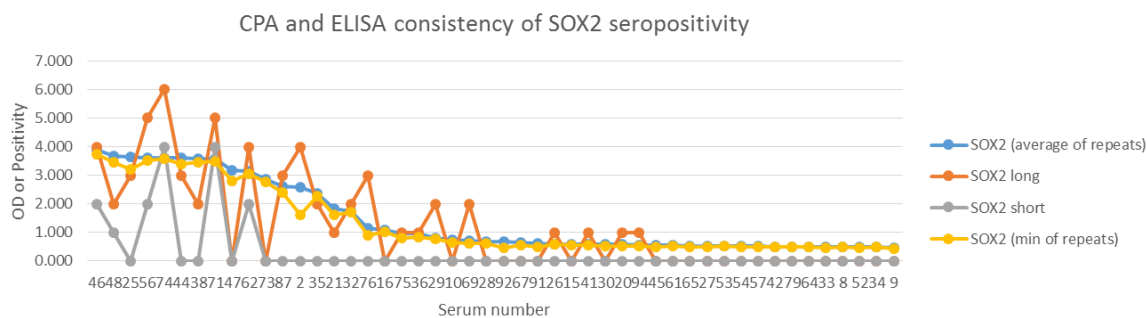


Figure 26. ELISA and Western blot correlation for a group of patients for the 6 antigens. 500 ng for each of the most valuable 6 antigens were run on SDS-PAGE and stained with Coomassie Brilliant blue G250 (Coomassie). The same order and amount for each protein was utilized for Western blot for a group of representative sera that are seropositive to at least one of the 6 antigens by ELISA. Corresponding ELISA values for each protein are shown below each Western blot.

ELISA and Western blot correlation is not only limited to SOX2, but also for the other valuable proteins that we have identified as the final panel of proteins after completing the ELISA screenings and their detailed analysis. These proteins are SOX2, NY-ESO-1, p53, C11orf20, ALMS1 and POLB. The upper part shows the Western blotting results and the lower values are the corresponding ELISA values. The higher the ELISA values the darker the shading (Figure 32).

5.3.12ELISA and custom PA comparison

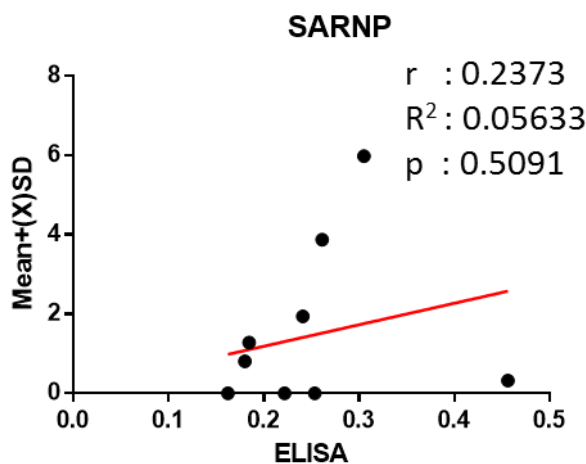
A)



B)

Parameter	SOX2 long	SOX2 short
Number of XY Pairs	48	48
Pearson r	0.7779	0.5905
95% confidence interval	0.6339 to 0.8698	0.3681 to 0.7490
P value (two-tailed)	< 0.0001	< 0.0001
P value summary	***	***
Is the correlation significant? (alpha=0.05)	Yes	Yes
R square	0.6052	0.3487

C)



D)

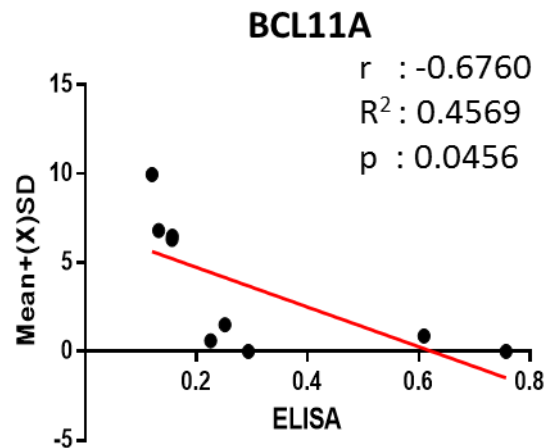
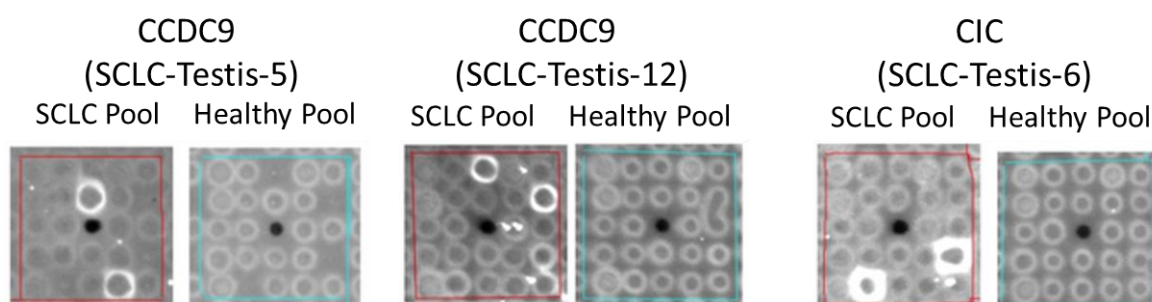


Figure 27. ELISA and custom PA correlation. **A)** Graph showing the SAS and ELISA correlation for SOX2. ELISA (average of repeats) values were sorted from highest to lowest for a group of patients. The correlation of these values with the minimum of repeats for ELISA and SAS values for two variants of SOX2 on custom PA (the long form and the short form) were shown to assess the level of correlation. **B)** Pearson's r correlation graph showing the significance of the correlations in A for the ELISA (average of repeats) values. **C)** Graph showing a positive correlation of PC values and ELISA values of a custom PA antigen; SARNP. **D)** Graph showing a negative correlation of PC values and ELISA values of a custom PA antigen; BCL11A.

Other than the ELISA and Western blot comparison we have also performed ELISA and custom PA comparison (Figure 33). We have observed a good correlation between ELISA values and their corresponding custom PA results (A). The graph shows this correlation for long and short variants of SOX2 (long variant includes 26-311 whereas short variant includes 164-311 of the 311 amino acids of SOX2). The table shows the Pearson's R values as a measure of correlation (B). There is significant correlation ($p < 0.0001$) both for the short and the long variant. The graphs in C) and D) show the correlation and its trend for other representative antigens. If outlier values are excluded there is a good correlation of ELISA with CPA for SARNP (C). However there is a negative correlation for B-7 which is unlikely (D).

5.3.13 Custom PA and Western blot comparison

A)



B)

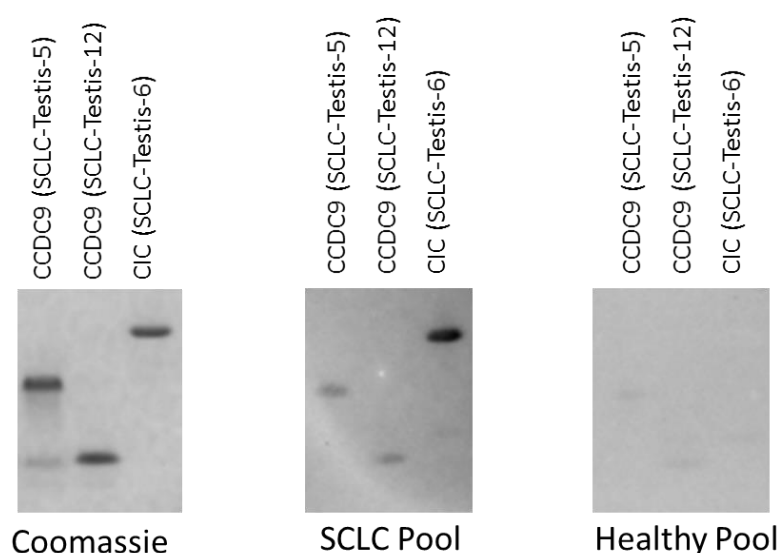


Figure 28. Correlation of PA with Western blotting for representative clones. A) PA pool screening results are shown both for SCLC pool and healthy pool for each of CCDC9 (SCLC-Testis-5), CCDC9 (SCLC-Testis-12) and CIC (SCLC-Testis-6), some of the strong positive signals by PA screening. **B)** Coomassie and Western blot results for the same proteins with pool screening are shown for comparison. 500 ng of pure protein was analyzed both for Coomassie and Western blot. 1:500 pool dilution ratio was applied both for PA screening and Western blot.

We have also validated our PA screening results with Western blotting (Figure 34). On the upper row the coomassie (SDS-PAGE), pool screening of SCLC and healthy screening results of CCDC9 and CIC are shown. On the lower row the PA screening results for these antigens are shown again for pool screening of SCLC and healthy controls. These

comparisons show that there is a high level of correlation between our Western blot and PA screening results.

5.3.14 Sensitivity and Specificity analysis

Table 18: Individual sensitivity values of the most valuable 6 antigens in 100%-90% specificity range by MC

Sensitivity (%)						Specificity (%)
SOX2	p53	POLB	C11orf20	NY-ESO-1	ALMS1	
14	2	2	2	2	1	100
14	7	4	5	2	2	99
16	9	6	7	2	3	98
18	10	11	8	4	4	97
19	11	13	9	5	7	96
19	12	16	10	5	7	95
20	13	17	12	5	8	94
21	15	17	13	5	9	93
22	16	19	14	5	10	92
22	20	19	16	6	10	91
22	20	21	18	7	10	90

After applying the different cut-off determination techniques and their evaluation we have decided to utilize Monte Carlo method since this method provides the calculation of the maximum sensitivity and specificity values for each antigen and for the panel by specifying the ideal cut-off values. The sensitivity values for each antigen are shown in this table for each specificity value between 100 and 90. The bigger the value the darker the shade of the cell. SOX2 has the highest sensitivity values when specificity reaches 100%, whereas both SOX2, p53, POLB and C11orf20 have similar sensitivity values when specificity reaches 90% (Table 18).

Table 19: Sensitivity and specificity values of the antigen panels for the best MC cut-offs

Sensitivity (%)				Specificity (%)		
All SCLC Cohorts	SCLC Cohort1	SCLC Cohort2	SCLC Cohort3	All Healthy Cohorts	Healthy Cohort1	Healthy Cohort2
18	28	17	11	100	100	100
21	28	21	14	99	100	99
23	30	23	19	98	100	98
26	32	26	20	97	100	97
29	40	28	21	96	100	96
31	48	27	24	95	100	94
33	51	30	27	94	97	94
35	51	30	32	93	97	92
38	53	33	34	92	94	92
39	54	34	36	91	94	90
40	54	35	39	90	94	89

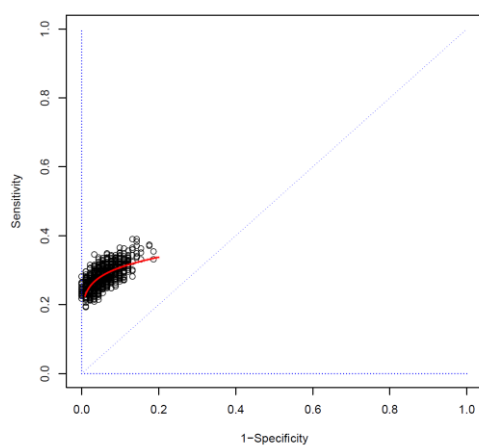
Monte Carlo method was used to determine cut-offs both to calculate individual sensitivity values of each antigen as well as panel sensitivity values. Table 19 shows the panel sensitivity values for the SCLC and healthy cohorts for the specificity values between 100 and 89. The bigger the value the darker the shade of the cell. As shown in the table SCLC Cohort1 has the highest sensitivity values and this is an expected situation since it is the cohort used to identify the autologous antibodies. SCLC Cohort2 and SCLC Cohort3 are validation cohorts and they have similar sensitivity values which is a good quality control of the analysis. Healthy cohorts also have similar specificity values which is another quality control measure.

5.3.15 ROC analysis

A)

SVM

SOX2 + p53

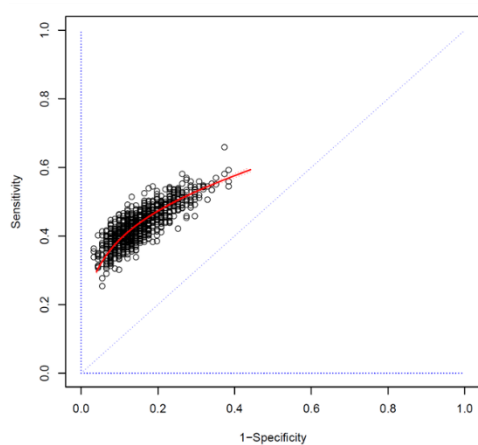


Sensitivity(%)	Specificity(%)
20	100
27	95
30	90

B)

SVM

SOX2 + p53 + POLB + C11orf20



Sensitivity(%)	Specificity(%)
20	100
31	95
38	90

C) MC

SOX2 + p53 + POLB + C11orf20

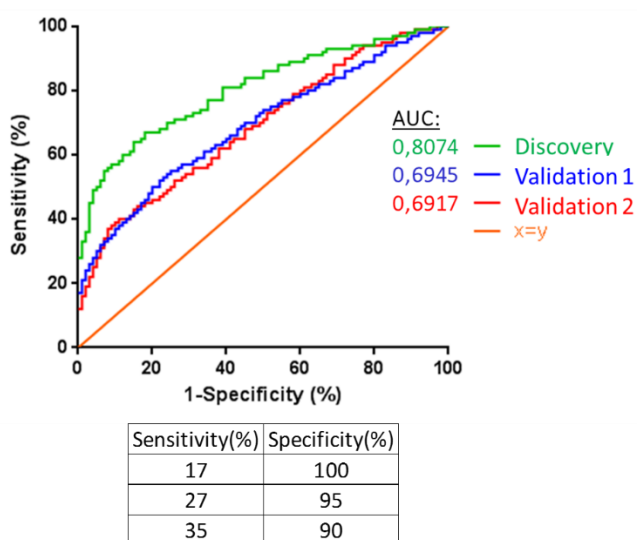


Figure 29. SVM and MC comparison for ELISA values. **A)** SVM was utilized by R for SOX2 and p53 and the resulting sensitivity and specificity values are plotted on the graph as black circles for SCLC Cohort2 (Validation 1). The red line is fitted to these values as an average ROC. The sensitivity and specificity values are also shown. **B)** SVM was utilized by R for SOX2, p53, POLB and C11orf20 and the resulting sensitivity and specificity values are plotted on the graph as black circles for SCLC Cohort2 (Validation 1). **C)** ROC curves obtained by MC cut-offs by SOX2, p53, POLB and C11orf20 both for the discovery (SCLC Cohort1) and validation cohorts (SCLC Cohort2, and SCLC Cohort3). Sensitivity and specificity values were shown for the SCLC Cohort2 (Validation 1).

We have analyzed the value of the panel of SOX2 and p53 by SVM algorithms by R for SCLC Cohort2 (Validation 1) and tried to identify the best panel sensitivity and specificity values and generated an ROC plot by these values (Figure 35, A). The values are indicated by circles and a red line indicating an ROC curve as the average of all data points. We have also utilized SVM algorithms for the panel of SOX2, p53, POLB and C11orf20 by R for SCLC Cohort2 (Validation 1) and tried to identify the best panel sensitivity and specificity values and generated an ROC plot by these values (Figure 35, B). The red line is very consistent with the panel ROC line generated with the Monte Carlo method (Figure 35, C). The average AUC of the ROC generated for the panel of SOX2 and p53 (Figure 35 A) is improved by addition of POLB and C11orf20 to the panel (Figure 35 B). Addition of these 2 antigens significantly

improved the ROC for the specificity values close to 100%. So, that means involvement of POLB and C11orf20 in the panel has a significant advantage. So that the panel of SOX2, p53, POLB and C11orf20 was identified as the best antigen panel for utilizing for discriminating SCLC from healthy controls.

The ROC lines are for the 3 SCLC cohorts we have screened with ELISA (Figure 35, C). As shown the ROC curve generated for the discovery cohort has a bigger AUC compared to the validation cohorts. This is an expected result since the discovery cohort was utilized to identify the autologous antibodies. The similarity between the ROC curves generated both by SVM and Monte Carlo indicates that the best panel sensitivity values are very consistent even when the evaluation method is changed.

5.3.16 Accuracy of antibody combination

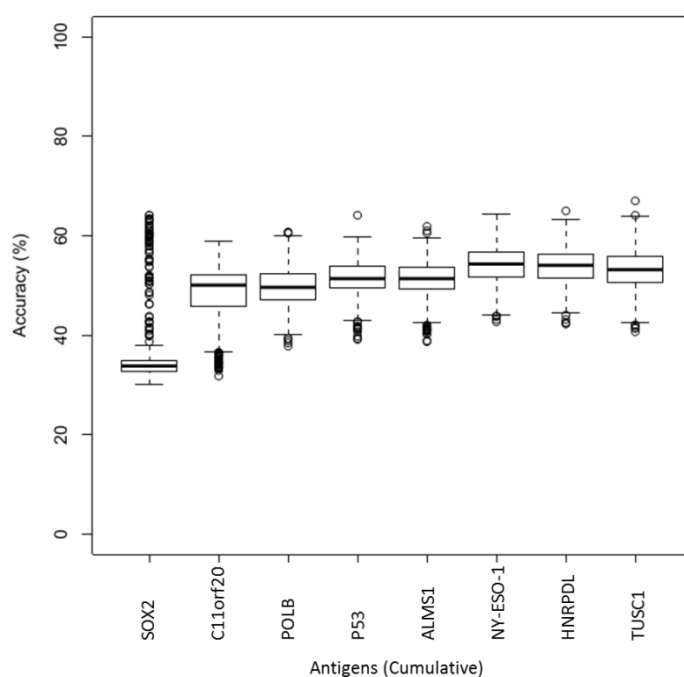


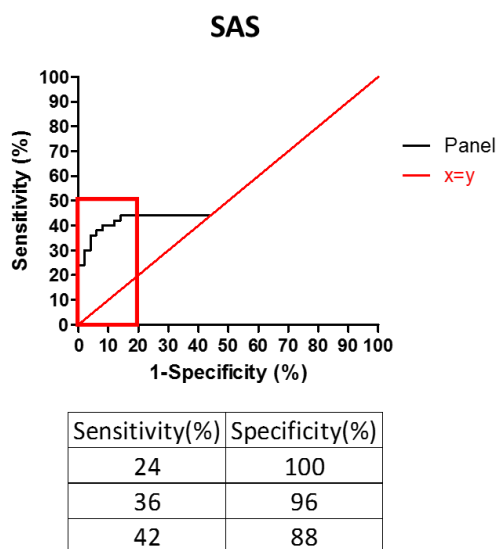
Figure 30. Feature selection for the antigen panel accuracy evaluation. Accuracy calculation was performed by R. Graph shows the accuracy values for cumulative inclusion of each antigen into the panel. Boxes indicate the 50 percentile while whiskers indicate 10-90 percentiles, horizontal bold lines indicate median values.

We have also performed a feature selection analysis by R. Accuracy measures how correct a diagnostic test identifies and excludes a given condition by calculating the ratio of $(TP+TN) / (\text{Number of all assessments})$. The graph shows the change in accuracy when each

autologous antibody indicated on the x axis is added in the evaluation. The accuracy analysis validates the most valuable antibody panel to be composed of SOX2, C11orf20, POLB and p53 (Figure 36).

5.3.17 Comparison of evaluation methods

A)



B)

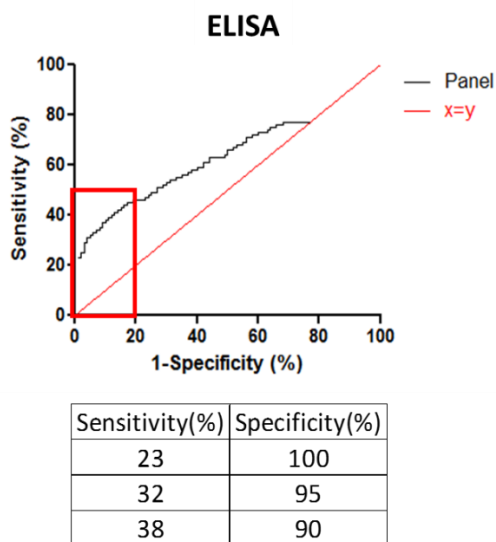


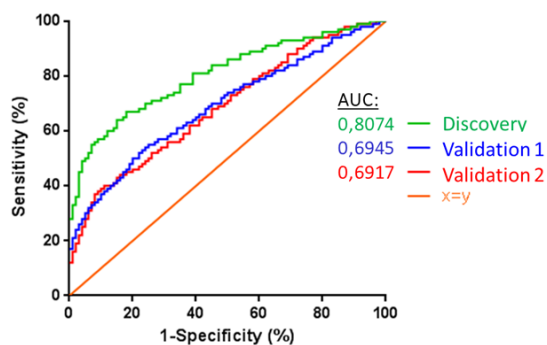
Figure 31. Comparison of SAS and ELISA for ROC by MC. **A)** MC was utilized to the SAS values of the discovery cohort for identification of the best panels. Sensitivity and specificity values of each panel was utilized for ROC plotting. Red line indicates $x=y$. **B)** MC was utilized to the ELISA values of the discovery cohort for identification of the best cut-offs for the most valuable 4 antigens. Sensitivity and specificity values of each cut-off was utilized for ROC plotting. Red line indicates $x=y$. Tables show some of the sensitivity and specificity values for comparison (**A** and **B**).

We have compared the ROC graphs plotted by the values generated by MC both for SAS (A) and ELISA (B) (Figure 37). For the 100%-80% specificity range the sensitivity values show a very similar trend. This result indicates even if the method of evaluation changes the discrimination power of the biomarkers doesn't change significantly. It is possible to reach a sensitivity value of 24% at 100% specificity by SAS whereas a sensitivity of 23% can be reached at 100% specificity by ELISA. The sensitivity can increase to 42% when the specificity drops to 88% for SAS values while the sensitivity increases to 38% when the specificity drops to 90% for ELISA values.

A)

Monte Carlo (ELISA)

SCLC Panel of
SOX2, p53, POLB, C11orf20

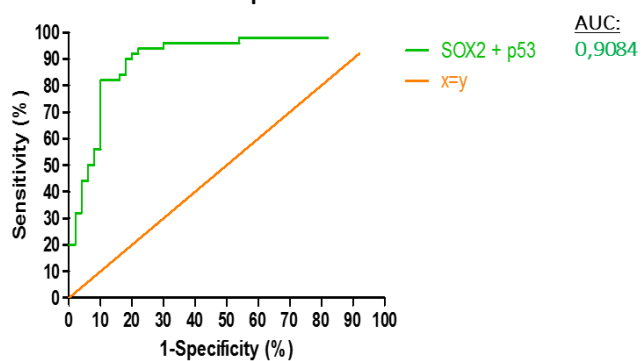


Sensitivity(%)	Specificity(%)
26	100
45	96
55	90

B)

Pixel Count (Custom PA)

SCLC Panel of
SOX2 and p53



Sensitivity(%)	Specificity(%)
20	100
44	96
82	90

C)

	Sensitivity(%)	Specificity(%)
Chest X-ray	65	96
Low-dose CT	89	92

Figure 32. ELISA, PC by MC and Chest X-ray and Low-dose CT comparison for ROC. A) ROC curves obtained by MC cut-offs for ELISA data of the panel of SOX2, p53, POLB and C11orf20 both for the discovery (SCLC Cohort1) and validation cohorts (SCLC Cohort2, and SCLC Cohort3). Sensitivity and specificity values were shown for the discovery cohort. **B)** ROC curve obtained for the discovery cohort only by utilization of SOX2 and p53 by the PC method (Original data generated by Alper Poyraz). **A, B)** Orange line indicates $x=y$. AUC values are indicated for each cohort. **C)** Sensitivity and specificity values for Chest X-ray and low-dose CT are shown for

We have compared ELISA (A) with PC (B) by MC cut-offs and chest X-ray and low-dose CT (C) in terms of sensitivity and specificity (Figure 38). Compared to ELISA values PC values are superior; AUC by PC for a panel only composed of SOX2 and p53 is 0.9084 whereas AUC for SOX2, p53, POLB and C11orf20 is 0.8074 for the discovery cohort, 0.6945 and 0.6917 for the validation cohorts. PC sensitivity and specificity values are comparable to that of low-dose CT so that have the potential to exceed the diagnostic power of both chest X-ray and low-dose CT if more antibodies are included in the panel.

5.4 PROGNOSTIC ANALYSIS OF AUTOLOGOUS ANTIBODIES

Table 20: Univariate Cox regression analysis for overall survival

SCLC Cohort2

Age	n	%	MS	HR	95%CI	p
<60	88	37	12.5	1		
>=60	150	63	9.5	1.399	1.049 - 1.865	0.022
Unknown	0	0				
Gender	n	%	MS	HR	95%CI	p
Female	40	17	11.1	1		
Male	198	83	10.2	1.087	0.758 - 1.559	0.649
Unknown	0	0				
Stage	n	%	MS	HR	95%CI	p
1	79	33	15	1		
02-Mar	159	67	8.4	2.291	1.683 - 3.119	<0.001
Unknown	0	0				
Performance Status	n	%	MS	HR	95%CI	p
0-1	153	64	13.8	1		
02-03-04	81	34	4.6	2.894	2.162 - 3.872	<0.001
Unknown	4	2				
Alkaline Phosphatase	n	%	MS	HR	95%CI	p
<90IU	114	48	11	1		
>=90IU	114	48	9.7	1.301	0.982 - 1.723	0.066
Unknown	10	4				
Lactate Dehydrogenase	n	%	MS	HR	95%CI	p
<275U	19	8	15	1		
>=275U	211	89	10.2	1.543	0.894 - 2.661	0.116
Unknown	8	3				

RespTher	n	%	MS	HR	95%CI	p
01-Feb	72	30	17.9	1		
Others	112	47	10.2	2.685	1.904 - 3.786	<0.001
Unknown	54	23				

SCLC Cohort3

Age	n	%	MS	HR	95%CI	p
<60	54	62	17.6	1		
>=60	31	36	10.9	1.724	0.997 - 2.984	0.049
Unknown	2	2				
Gender	n	%	MS	HR	95%CI	p
Female	10	11	13.1	1		
Male	77	89	13.8	1.195	0.512 - 2.793	0.68
Unknown	0	0				
Stage	n	%	MS	HR	95%CI	p
1	40	46	20.3	1		
02-Mar	47	54	10.8	2.451	1.432 - 4.193	0.001
Unknown	0	0				
Performance Status	n	%	MS	HR	95%CI	p
0-1	64	74	17.4	1		
02-03-04	12	14	9.8	2.299	1.048 - 5.043	0.033
Unknown	11	13				
Alkaline Phosphatase	n	%	MS	HR	95%CI	p
<90IU	48	55	13.1	1		
>=90IU	39	45	17.8	0.855	0.493 - 1.48	0.575
Unknown	0	0				
Lactate Dehydrogenase	n	%	MS	HR	95%CI	p
<275U	12	14	25.2	1		
>=275U	65	75	12.3	2.749	1.203 - 6.286	0.013
Unknown	10	11				
RespTher	n	%	MS	HR	95%CI	p
01-Feb	66	76	17.6	1		

Others	16	18	4.2	10.698	4.841 - 23.644	<0.001
Unknown	5	6				

Univariate Cox Regression analysis was performed to the 2 validation cohorts that clinical data was available (Table 20). Also these 2 cohorts showed a very similar ROC from the ELISA values by Monte Carlo. This analysis was performed to analyze the correlation between clinical parameters with overall survival. The statistically significant parameters are indicated with bold characters. The comparison of the 2 cohorts show that except lactate dehydrogenase levels all of age, stage, performance status and response to therapy are statistically significantly affect overall survival for both cohorts.

Table 21: Antibody frequency and overall survival (cut-off: Mean + 3 * StDev)

SCLC Cohort2

SOX2	n	%	MS	HR	95%CI	p
<M+15SD	202	85	9.9	1		
>=M+15SD	31	13	14.9	0.559	0.335 - 0.933	0.024
Unknown	5	2				
P53	n	%	MS	HR	95%CI	p
<M+3SD	219	92	10.5	1		
>=M+3SD	15	6	10.2	1.052	0.599 - 1.848	0.86
Unknown	4	2				
POLB	n	%	MS	HR	95%CI	p
<M+3SD	225	95	10.3	1		
>=M+3SD	9	4	14.9	0.692	0.341 - 1.408	0.307
Unknown	4	2				
C11orf20	n	%	MS	HR	95%CI	p
<M+3SD	225	95	10.5	1		
>=M+3SD	9	4	12.6	0.88	0.45 - 1.72	0.708
Unknown	4	2				
TUSC1	n	%	MS	HR	95%CI	p
<M+3SD	229	96	10.3	1		
>=M+3SD	4	2	21.8	0.243	0.06 - 0.982	0.031
Unknown	5	2				
ALMS1	n	%	MS	HR	95%CI	p

<M+3SD	232	97	10.5	1		
>=M+3SD	2	1	21.8	0.229	0.032 - 1.638	0.109
Unknown	4	2				
NY.ESO.1	n	%	MS	HR	95%CI	p
<M+3SD	226	95	10.5	1		
>=M+3SD	8	3	11.2	0.915	0.405 - 2.067	0.831
Unknown	4	2				
HNRPDL	n	%	MS	HR	95%CI	p
<M+3SD	232	97	10.6	1		
>=M+3SD	2	1	5.3	3.141	0.769 - 12.82	0.092
Unknown	4	2				

SCLC Cohort3

SOX2	n	%	MS	HR	95%CI	p
<M+3SD	72	83	13.8	1		
>=M+3SD	13	15	12.3	0.948	0.445 - 2.02	0.891
Unknown	2	2				
P53	n	%	MS	HR	95%CI	p
<M+3SD	82	94	16.6	1		
>=M+3SD	3	3	13.8	2.242	0.536 - 9.375	0.256
Unknown	2	2				
POLB	n	%	MS	HR	95%CI	p
<M+3SD	82	94	13.8	1		
>=M+3SD	3	3	10.8	1.33	0.413 - 4.285	0.632
Unknown	2	2				
C11orf20	n	%	MS	HR	95%CI	p
<M+3SD	84	97	13.8	1		
>=M+3SD	1	1	NA	0	0 - Inf	0.736
Unknown	2	2				
TUSC1	n	%	MS	HR	95%CI	p
<M+3SD	83	95	16.6	1		
>=M+3SD	1	1	7.9	4.577	0.597 - 35.074	0.108
Unknown	3	3				
ALMS1	n	%	MS	HR	95%CI	p
<M+3SD	79	91	17.1	1		
>=M+3SD	5	6	8.7	3.613	1.238 - 10.548	0.012

Unknown	3	3				
NY.ESO.1	n	%	MS	HR	95%CI	p
<M+3SD	80	92	13.8	1		
>=M+3SD	4	5	14.5	1.597	0.488 - 5.22	0.434
Unknown	3	3				
HNRPD L	n	%	MS	HR	95%CI	p
<M+3SD	83	95	13.8	1		
>=M+3SD	1	1	NA	0	0 - Inf	0.734
Unknown	3	3				

We have also analyzed the correlation of antibody positivity with overall survival for both cohorts (Table 21). However there was no consensus for antibody positivity in both cohorts. SOX2 and TUSC1 autologous antibodies were statistically significantly correlating with good prognosis for SCLC Cohort2 whereas ALMS1 autologous antibodies were statistically significantly correlating with poor prognosis for SCLC Cohort3, and the other autologous antibody responses were not correlating with overall survival. This situation indicates different cohorts have distinct autologous antibody signatures and at least in our case it is not possible to identify a common autologous antibody response that is statistically significantly correlating with survival in multiple cohorts, most probably due to SCLC tumor heterogeneity.

Table 22: Cox Regression analysis and overall survival

A) SCLC Cohort2

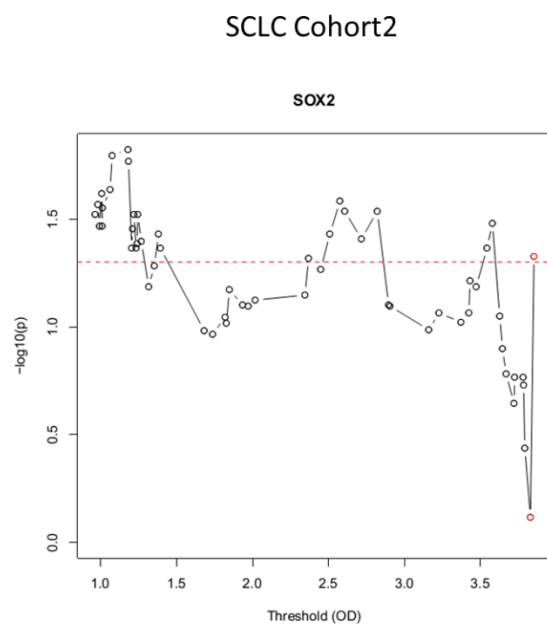
		B	SE	Wald	df	Sig.	Exp(B)	95.0% CI for Exp(B)	
								Lower	Upper
Step 1	SOX2	-.190	.084	5.086	1	.024	.827	.702	.975
	p53	-.015	.139	.012	1	.913	.985	.750	1.294
	C11orf20	-.136	.162	.708	1	.400	.873	.635	1.199
	ALMS1	-.140	.695	.040	1	.841	.870	.223	3.398
	TUSC1	-1.769	.868	4.155	1	.042	.170	.031	.934
	POLB	-.119	.749	.025	1	.874	1.126	.259	4.894
	NY-ESO-1	-.183	.165	1.219	1	.270	1.200	.868	1.660
	HNRPD L	-.531	.328	2.631	1	.105	1.701	.895	3.233

B) SCLC Cohort3

		B	SE	Wald	df	Sig.	Exp(B)	95.0% CI for Exp(B)	
								Lower	Upper
Step 1	SOX2	-.013	.171	.006	1	.939	.987	.706	1.380
	p53	.078	.735	.011	1	.915	1.082	.256	4.566
	C11orf20	.496	.399	1.540	1	.215	1.642	.750	3.591
	ALMS1	-.899	.989	.827	1	.363	.407	.059	2.826
	TUSC1	.745	1.188	.393	1	.531	2.106	.205	21.602
	POLB	-.144	1.187	.015	1	.904	.866	.085	8.879
	NY-ESO-1	-.161	.441	.134	1	.714	.851	.359	2.018
	HNRPD	.278	.835	.110	1	.740	1.320	.257	6.785

We have also performed multivariate cox regression analysis with the autologous antibody responses individually statistically significantly correlating with overall survival to analyze their combined effect on overall survival (Table 22). SOX2 and TUSC1 autologous antibody responses are statistically significantly correlating with good overall survival when all the autologous antibodies are evaluated together for SCLC Cohort2. However this significance disappears when only these two autologous antibody responses are analyzed with multivariate cox regression analysis for overall survival. For SCLC Cohort3 combined evaluation of any of the autologous antibody responses doesn't correlate with overall survival significantly. ALMS1 autologous antibodies correlated with worse overall survival.

A)



B)

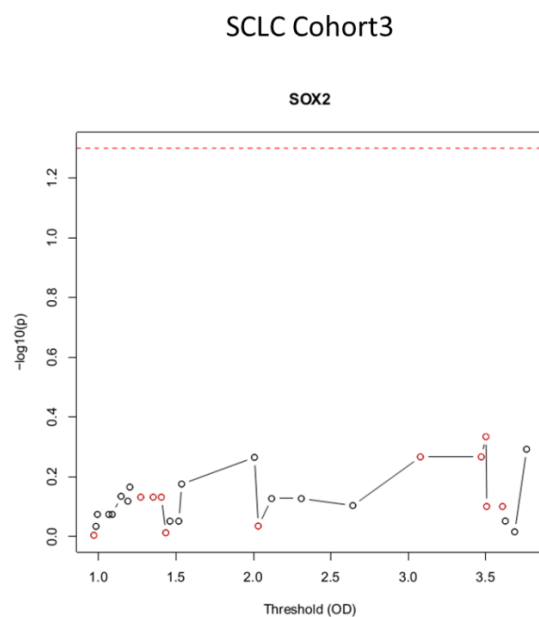
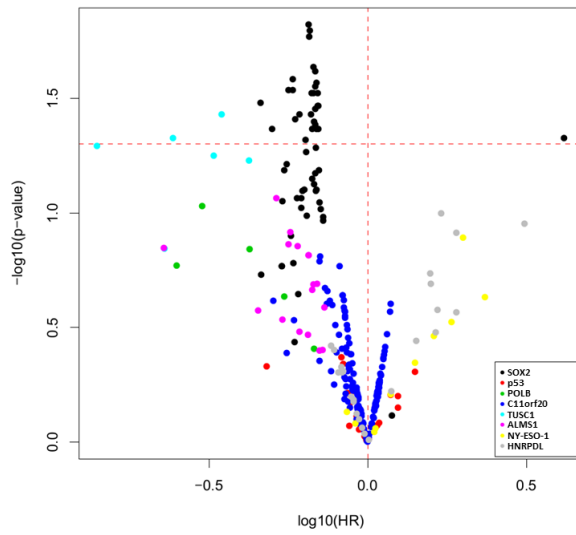


Figure 33. Seropositivity and survival correlation for different cut-offs. Each cut-off was analyzed for correlation with good or worse prognosis. Good prognosis is indicated by black circles while worse prognosis is indicated by red circles. Statistical significance was indicated with the horizontal red dashed line. Significance cut-off was $p=0.05$. The analysis was performed both for **A)** SCLC Cohort2 and **B)** SCLC Cohort3 for which clinical data is available.

Other than the Cox Regression analysis we have also performed log-rank analysis for evaluation of different cut-offs for the significance of the autologous antibody positivity correlation with overall survival for the autologous antibodies that were significantly correlating with overall survival (Figure 39). These graphics show the significance of different cut-offs for worse (black circles) and good (red circles) prognosis. These graphs show the result of the analysis for SOX2 both in SCLC Cohort2 and SCLC Cohort3 and the red dashed horizontal lines show the significance cut-off. SOX2 autologous antibodies are statistically significantly correlating with worse prognosis for the cut-offs between 1.0 to 1.5 for SCLC Cohort2. Whereas any cut-off is not statistically significant for correlation of SOX2 autologous antibodies with good or worse prognosis for SCLC Cohort3.

A)

SCLC Cohort2



B)

SCLC Cohort3

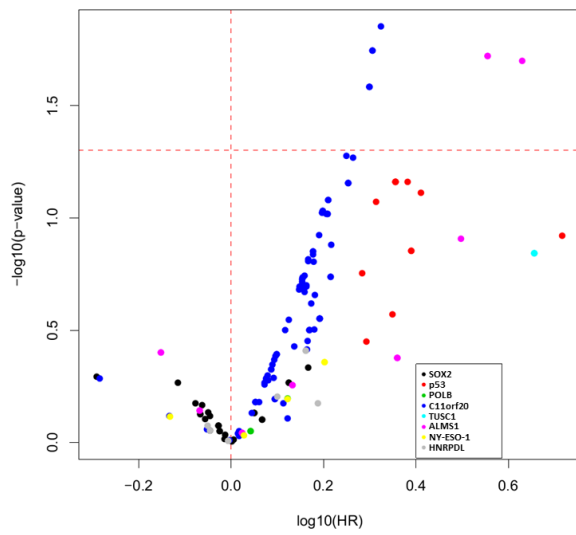
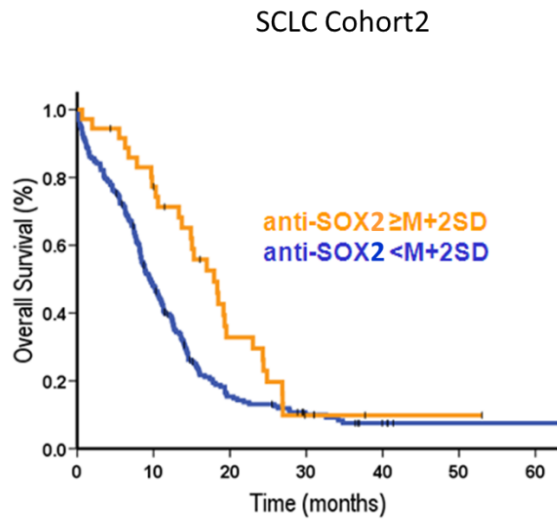


Figure 34. Volcano plots; seropositivity and survival correlation for SCLC cohorts. 8 ELISA antigens were analyzed for correlation with survival at different cut-offs. Right side of the vertical red dashed line corresponds to correlation with worse prognosis whereas left side with good prognosis. Horizontal dashed red line indicates significance cut-off, $p=0.05$. Each antigen is represented with a different color and the cut-offs above $M+1\text{StDev}$ are shown. The graphs are plotted for **A)** SCLC Cohort2 and **B)** SCLC Cohort3 for which clinical data is available.

We have also formed Volcano plots for showing the cut-offs above $M+1\text{SD}$ for correlation with survival and its significance for all 8 ELISA antigens that we have complete ELISA data (Figure 40). The horizontal dashed red line is the significance cut-off and the dots on the left of the vertical dashed red line means at that cut-off the antibody response correlates with overall survival whereas for the dots on the right it correlates with worse prognosis. Each autologous antibody was represented with a different color and obviously these 2 cohorts have a different profile in terms of autologous antibody seropositivity and survival correlation. For example SOX2 antibody response is significantly correlating with good prognosis for SCLC Cohort2 whereas it doesn't correlate with good prognosis for SCLC Cohort3 even though there is a trend. This situation indicates that for different cohorts autologous antibody response doesn't show a common behavior in the context of prognosis.

A)



B)

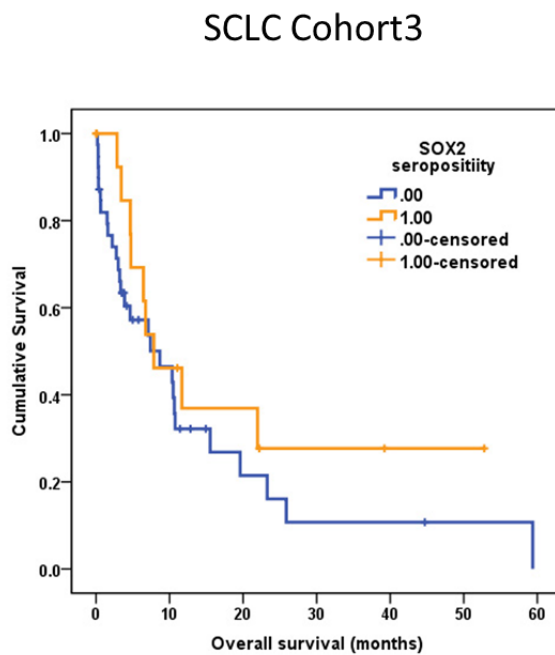


Figure 35. Seropositivity and overall survival correlation by Kaplan-Meier graphs. SOX2 seropositivity was analyzed for survival correlation for the M+2StDev cut-off. Kaplan-Meier curves were plotted both for **A)** SCLC Cohort2 and **B)** SCLC Cohort3.

The correlation between seropositivity and prognosis is also observable and consistent with the Kaplan-Meier analysis. SOX2 antibody seropositivity according to the M+2SD cut-

off correlate with overall survival for SCLC Cohort2 whereas there is no significant correlation for SCLC Cohort3 (Figure 41). An independent study also shows that there is no survival advantage for the group of SCLC patients who can elicit SOX1 autologous antibody response in their cohort [70].

SOX1 and SOX2 proteins have a high level of amino acid sequence homology (>90%) and in >90% of the cases the autologous antibodies elicited to either of the proteins recognize the other. These results indicate that SOX2 antibody seropositivity doesn't show a consensus as a prognostic marker of overall survival for different cohorts, most probably due to tumor heterogeneity and thus different cohorts are not comparable.

5.5 ANTIGEN EXPRESSION AND ANTIBODY CORRELATION

A)

		n	Median survival (months)	P (log rank)
Age	61 (44–78)			
	<60	24	10.36	0.76
	≥60	35	7.1	
LDH	193 IU/L (133–380)			
	<200	30	10.75	0.36
	≥200	25	8.86	
	U*	4		
AP	93 IU/L (50–171)			
	<70	14	3.97	0.47
	≥70	44	8.68	
	U	1		
Stage	Limited	28	10.36	0.03
	Extensive	27	7.34	
	U	1		
SOX2 antibody	<M + 2SD	42	8.68	0.3
	≥ M + 2SD	17	7.85	
SOX2 IHC intensity**	0-1	25	7.39	0.98
	2-3	30	10.65	
SOX2 IHC frequency	<5 %	16	25.87	0.12
	≥5 %	39	7.39	
SOX2 IHC frequency	<20 %	24	7.4	0.11
	≥20 %	31	7.86	
SOX2 IHC frequency	<40 %	29	7.39	0.43
	≥40 %	26	9.68	

*U: unknown. **Tumor tissue sections evaluations were available from 55 patients.

B)

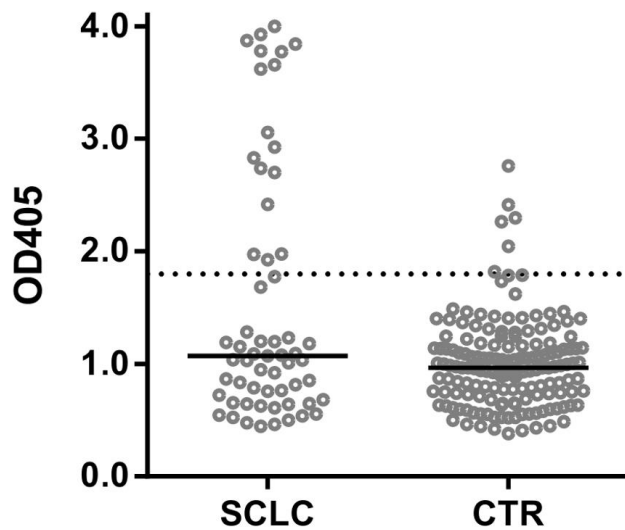


Figure 36. SOX2 autologous antibody and SOX2 expression correlation. A) Univariate Cox regression analysis of clinical features; age, LDH, AP, stage, SOX2 antibody, SOX2 IHC intensity and frequency for different cut-offs of the 59 SCLC patients and their association with overall survival. Significance was indicated as bold. **B)** Scatter dot plots of the ELISA values of 59 SCLC and 192 healthy controls. Horizontal bars show median values for both groups and horizontal dashed line indicates seropositivity cut-off.

In a separate study we have shown that the autologous anti-SOX2 antibody responses reflect intensity but not frequency of antigen expression in SCLC. In this study in a SCLC cohort of 59 patients and 192 healthy controls we have measured the SOX2 autologous antibody levels by ELISA as we have done before (Figure 42, A). The scatter dot plot shows the anti-SOX2 antibody levels of these 2 cohorts, the median values were indicated with the horizontal lines and the cut-off was shown with the dotted line (Figure 42, B). When analyzed for correlation with overall survival only stage was significant ($p=0.03$) among all the known clinical parameters together with SOX2 seropositivity. Neither SOX2 antigen expression intensity nor frequency were correlating with overall survival.

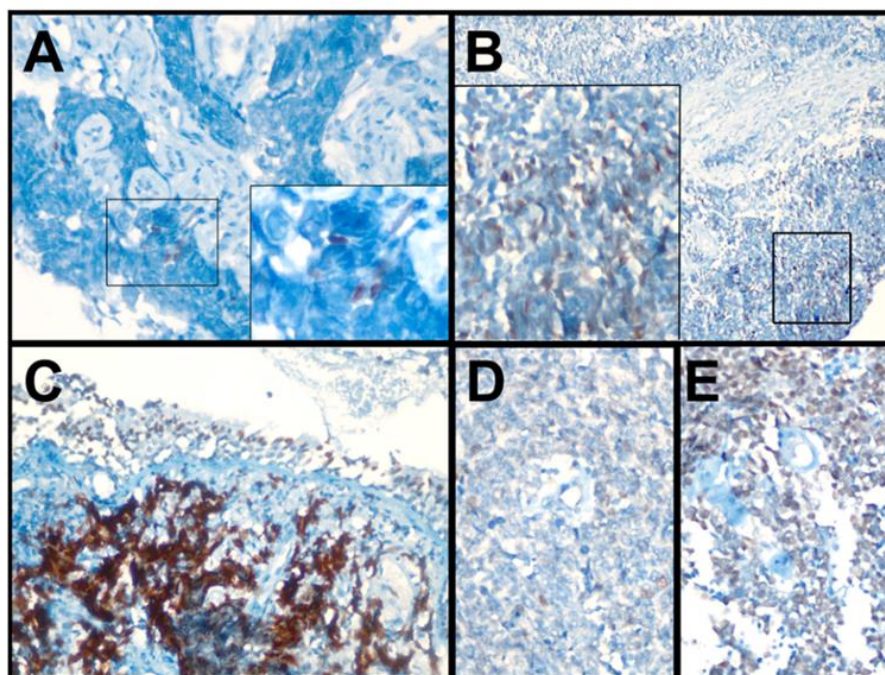
A)

		SOX2 Ab		P (chi.sq.)
		(-)	(+)	
Stage	Limited	16	12	0.05
	Extensive	22	5	

B)

		SOX2 Ab		P (chi.sq.)
		(-)	(+)	
SOX2 IHC intensity	0-1	23	2	0.017
	2-3	16	14	

C)



D)

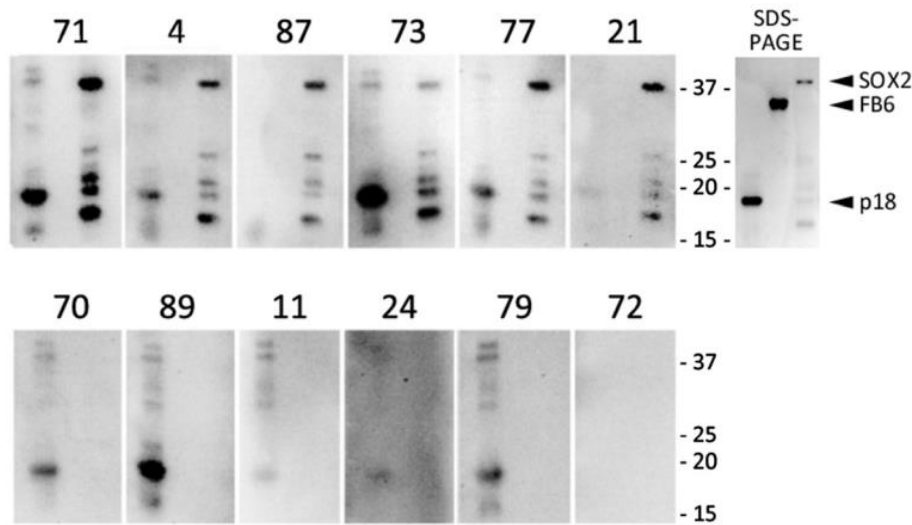


Figure 37. SOX2 protein expression and autologous antibody correlation analysis. SOX2 antibody correlation with clinical stage (A), and intensity of SOX2 protein expression (B). Correlation of SOX2 autologous antibodies with SOX2 protein expression intensity (C) Windows of A-E shows distinct IHC stainings of patient tissue samples against SOX2. Selected areas are magnified and shown in A and B windows. Confirmation of ELISA data with Western blot (D) Upper row shows Western blot results for SOX2 seropositive sera whereas lower row represents the seronegatives. p18 is the positive control. (Original data from A, B and C generated by Alper Poyraz)

We have shown that SOX2 autologous antibodies correlated with limited clinical stage significantly (Figure 43, A). Also elucidation of SOX2 autologous antibodies correlated significantly with SOX2 protein expression intensity (B). The IHC SOX2 stainings in the A, B and C windows shows intense SOX2 staining whereas D and E show weak intensity (C). In the study we were able to validate our ELISA results with Western blotting (D). The upper row includes SOX2 seropositive patients (according to ELISA data) and the lower images for SOX2 seronegative patients.

5.6 AUTOLOGOUS ANTIBODIES IN AUTOIMMUNE DISEASE

5.6.1 Autologous antibody identification for NBD

In a separate study our optimized protocols were utilized for identification of autologous antibodies for NBD. Mtch1 autologous antibodies were identified by screening of PA with NBD patients' sera in parallel with healthy controls. Mtch1 is an apoptosis related protein and it is highly probable that the autologous antibodies are relevant in terms of pathogenicity of NBD. According to the ELISA data Mtch1 antibodies were present in 68 out of 144 BD sera and 4 out of 168 control sera. These numbers corresponds to 47% sensitivity at 98% specificity. Mtch1 antibodies correlated with more attacks, higher level of disability and lower serum nucleosome levels.

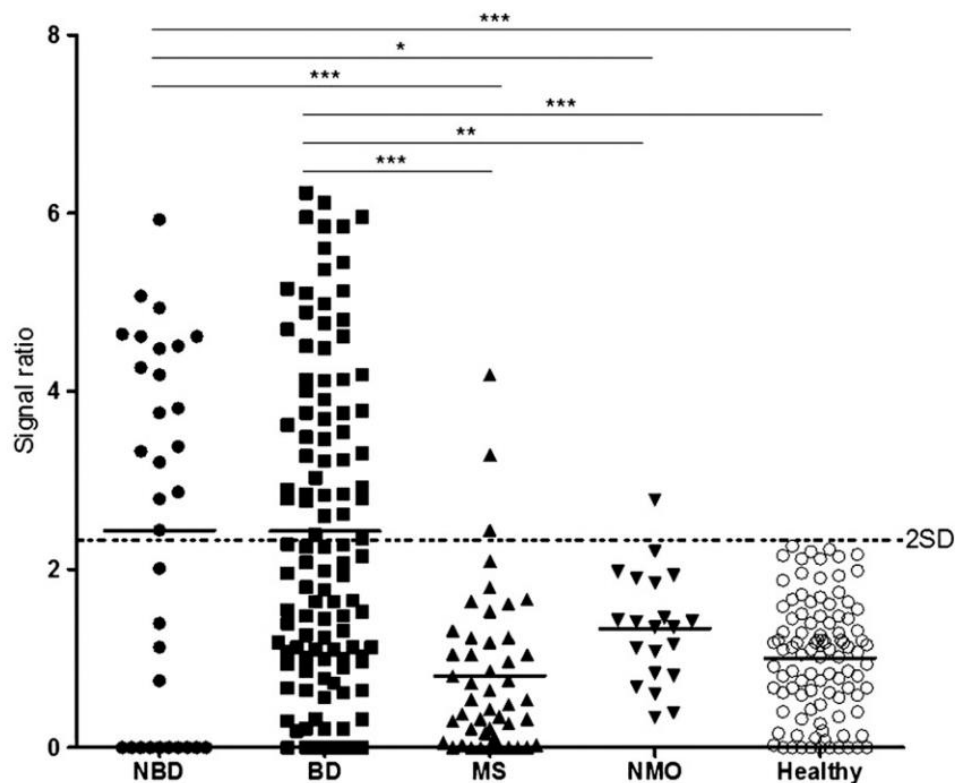


Figure 38. Scatter dot plots of Mtch1 antibodies in the cohorts. Horizontal dashed line indicates the seropositivity cut-off: $M+2StDev$ (calculated by healthy control values). Patient groups were NBD, BD (with no neurological involvement), multiple sclerosis (MS) and neuromyelitis optica (NMO). Horizontal bars represent mean values of each group.

ELISA values were used to generate the scatter dot plots for NBD, BD, multiple sclerosis (MS), neuromyelitis optica (NMO) patients and healthy controls (Figure 44). Compared to control groups, NBD and BD patients had significantly higher ELISA values as ($p < 0.0001$ by ANOVA and $p < 0.05$ – 0.001 by Tukey's post-hoc test).

Additionally pooled PA screening was also performed both with Testis and FB for identification of autologous antibodies specific to BD by comparison with the screenings performed with healthy serum pool. We have identified the most valuable clones from the pool screening and included them in custom PA spotting. After evaluation of the custom PA screenings the most valuable 8 clones were identified for BD (Table 23).

Table 23: Final list of valuable clones after custom PA screening of BD

No	Clone	Gene info	Protein Identity of clone
1	Behcet-FB-1	FAM32A (NM_014077.2)	112 aa (1-112) of the 112 aa
2	Behcet-FB-2	GON4L (NM_001037533.1)	205 aa (1953-2157) of the 2240 aa (24,7 kDa)
3	Behcet-FB-4	ANAPC5 (NM_016237.4)	681 aa (75-755) of the 755 aa
4	Behcet-FB-5	SARNP (NM_033082.3)	210 aa (1-210) of the 210 aa
5	Behcet-FB-7	BCL11A (NM_138559.1)	150 aa (93-243) of the 243 aa (19,4 kDa)
6	Behcet-FB-10	FTH1 (NM_002032.2)	153 aa (31-183) of the 183 aa
7	Behcet-FB-12	STMN4 (NM_030795.2)	189 aa (28-216) of the 216 aa (21,8 kDa)
8	Behcet-FB-13	CCNL2 (NM_030937.4)	255 aa (266-520) of the 520 aa (31,7 kDa)

The most valuable 8 clones were analyzed by MC for the SAS values for identification of the best panels. The most valuable panel was formed by 4 autologous antibodies; FAM32A, GON4L, ANAPC5 and SARNP and reached a sensitivity of 42% at 100% specificity (Table 24). Sensitivity value increased to 78% at 86% specificity by a panel of 7 antibodies. Further validation of these clones by ELISA and Western blot is going on as a separate study.

Table 24: Most valuable MC panels for BD and their antibody content

Panels	Sensitivity (%)	Specificity (%)	# of antigens	Behcet-FB-1	Behcet-FB-2	Behcet-FB-4	Behcet-FB-5	Behcet-FB-7	Behcet-FB-10	Behcet-FB-12	Behcet-FB-13	Clones
				FAM32A	GON4L	ANAPC5	SARNP	BCL11A	FTH1	STMN4	CCNL2	Genes
1	42	100	4	+	+	+	+	-	-	-	-	
2	60	98	5	+	+	+	+	+	-	-	-	
3	62	96	6	+	+	+	+	+	+	-	-	
4	48	88	5	+	+	+	+	-	-	+	-	
5	78	86	7	+	+	+	+	+	+	-	+	
6	68	84	7	+	+	+	+	+	+	+	-	
7	84	76	8	+	+	+	+	+	+	+	+	

Chapter 6

6 SUMMARY

In order to identify novel autologous anti-tumor antibodies specific to cancer we have developed and optimized a PA screening based method. We have screened commercial PAs from Imagenes by pooled serum of SCLC, colorectal, gastric, ovarian cancer, Behçet's disease and Healthy controls separately. Comparative analysis of the signals resulted in identification of cancer specific signals and we have further validated these signals for SCLC by ELISA. We were able to identify a panel of autologous anti-tumor antibodies valuable for SCLC diagnosis. We were also able to validate the antibody panel by Western blotting. We observed protein expression intensity and antibody presence correlation. We have also validated the utility of our method for identification and validation of autoimmune disease specific autologous antibodies.

6.1 Significance of the study

The study has multiple significant aspects since a number of innovations and comprehensive evaluations were performed throughout the study.

(1) A new method was developed for PA screening in which stable and constant signal intensity assured reliable, reproducible and comparable results which enabled comparison of cancer/disease and healthy controls, so that identification of cancer/disease specific signals became possible.

(2) We screened PAs with the biggest number of protein diversity and clone content to our knowledge so that a comprehensive autologous antibody identification methodology was achieved. Rearranging opportunity for the selected clones enabled precise sensitivity and specificity evaluation of the selected clones.

(3) The study was performed for multiple cancers. Comparative analysis revealed the method's utilization potential for identification of autologous antibody biomarkers specific for SCLC, ovarian, gastric, colorectal cancers and Behçet's disease.

(4) The final list of clones defined after PA screening evaluations were validated by both ELISA and Western blotting. This study is valuable since 3 distinct SCLC and 2 Healthy cohorts were screened by ELISA for validation.

(5) The total number of SCLC patients screened were close to 450 and the total number of healthy controls screened was close to 200, which increases the significance of the study. Prognostic correlation with antibody seropositivity was also performed for two of the SCLC cohorts that clinical data was available.

(6) We have also investigated the correlation between antibody seropositivity and antigen expression for SOX2 which was identified as the most valuable autologous antibody specific to SCLC. SOX2 antibody seropositivity was identified to be correlating with SOX2 antigen expression intensity which contributed explanation of antibody elucidation mechanisms.

(7) We have also shown that the autologous antibody identification and validation strategy can be applied for evaluation of autoimmune disease specific antibodies. The methodology that we have optimized was successfully applied for identification and validation of Mth1 autologous antibody as specific to NBD.

6.2 Major findings

We have identified 56 clones from Testis and 22 clones from FB PAs as SCLC specific. 27 of 56 and 15 of 22 were significantly stronger than the signals obtained from the healthy screenings. For both Testis and FB PA screenings there were both cancer specific and common cancer autologous antibodies were identified.

We have reidentified the already known autologous anti-tumor antibodies by PA screening. We have identified variants of the same antigen on the same and on different PAs. We observed stronger signal intensities from the cancer PA screenings compared to the screenings with the healthy controls, which indicates that we have identified cancer specific autologous antibodies. The clones selected as the final list were confirmed to be expressing a unique antigen in the correct open reading frame.

Respotted custom PAs were screened with individual sera for sensitivity and specificity calculations. MSS and SAS were utilized for signal quantification and MC was performed for ideal cut-off and panel identification. Resulting ROC curves showed the

diagnostic power of panels for each disease. For SCLC one of the best antibody panels had a sensitivity of 40% at 92% specificity with 8 antigens. For colorectal cancer a sensitivity of 44% was reached at 92% specificity with 8 antigens. A sensitivity of 39% is reached at 92% specificity with 6 antigens for gastric cancer. For Ovarian cancer 48% sensitivity is obtained at 92% specificity by an antigen panel of 10.

These clones selected from Testis PA screening were analyzed for their insert and plasmid content by restriction enzyme digestion and induced with IPTG for their protein expression capacity.

We have identified 23 SCLC clones for ELISA validation which have the highest value based on the custom PA screening evaluations and Testis PA pool screenings.

We have designed an iterative ELISA approach to screen the maximum number of valuable proteins with all SCLC cohorts in the least amount of time. We have repeatedly selected the most valuable antigens at each ELISA cycle by evaluation based on the Mean + 10, 7, 5, 4 and 3 StDev cut-offs. We have applied 2 levels of normalization to the ELISA data to obtain comparability and standardization of the data.

We have evaluated our ELISA data by calculating the CV for both high signals and medium signals. We have obtained higher CV values compared to EarlyCDT-Lung data which is the only diagnostic kit currently used in clinics. We have analyzed the antibody integrity level of the sera in our cohorts to eliminate the possibility of biased screening and false positives by screening the sera for EBV p18 seropositivity. We have observed no significant difference by t-test for the SCLC and Healthy cohorts as well as validation cohorts in terms of p18 seropositivity which indicates the sera in the cohorts are intact.

We have complete ELISA data for the most valuable antigens as an output of the iterative ELISA approach. We have utilized manual cut-off identification and MC for defining the ideal cut-offs for seropositivity for each antigen both individually and for the panel.

There is a significant correlation between our ELISA results and Western blot results, our ELISA results and custom PA results and our custom PA results and Western blot results, which indicates the ELISA data is reliable and reproducible.

We have obtained individual sensitivity values of 22%, 20%, 21% and 18% at 90% specificity for SOX2, p53, POLB and C11orf20 respectively. We have also obtained 54%,

35% and 39% panel sensitivity values at 90% specificity for our discovery and validation cohorts respectively. We have confirmed the accuracy of the most valuable antigen panel to be composed of SOX2, p53, POLB and C11orf20 by feature selection utilizing R.

We have compared the ROC of the panel composed of SOX2, p53, POLB and C11orf20 and the panel composed of SOX2 and p53 by SVM and revealed that addition of POLB and C11orf20 to the panel of SOX2 and p53 improved the AUC of the ROC. We have analyzed the ELISA data both by SVM and MC for ideal cut-off determination. The resulting ROC curves were very consistent with each other meaning both methods identify similar discriminative characteristics.

We have evaluated our ELISA results by MC to define the most valuable antigens to be in the panel and to define the ideal cut-offs for their combined evaluation. Custom PA results were also evaluated by MC for identification of the most valuable panels. ROC curves were plotted based on these analysis and both ELISA and custom PA data resulted in similar ROC curves for the 100%-80% specificity range. This data suggests that the discriminative power of the antibodies doesn't change significantly even if the method of evaluation changes.

We have developed a pixel counting (image processing) method called PC for evaluation of the custom PA results. PC values are superior over ELISA; AUC by PC for a panel only composed of SOX2 and p53 is 0.9084 whereas AUC by ELISA for a panel of SOX2, p53, POLB and C11orf20 is 0.8074 for the discovery cohort. PC sensitivity and specificity values are comparable to that of low-dose CT.

We have performed univariate and multivariate Cox regression analysis both with the clinical parameters and antibody seropositivities for the validation cohorts. For the univariate Cox regression analysis comparison of the 2 cohorts show that except lactate dehydrogenase levels all of age, stage, performance status and response to therapy are statistically significantly affect overall survival for both cohorts. For the seropositivity SOX2 and TUSC1 autologous antibodies were statistically significantly correlating with good prognosis for SCLC Cohort2 whereas ALMS1 autologous antibodies were statistically significantly correlating with poor prognosis for SCLC Cohort3. However there was no common correlation for both cohorts. For multivariate Cox regression analysis SOX2 and TUSC1 autologous antibody responses are statistically significantly correlating with good overall

survival when all the autologous antibodies are evaluated together for SCLC Cohort2. There was no correlation identified for SCLC Cohort3.

We have analyzed correlation of autologous antibody response at multiple seropositivity cut-offs and overall survival. SOX2 autologous antibodies are statistically significantly correlating with worse prognosis for the cut-offs between 1.0 to 1.5 for SCLC Cohort2. Whereas there was no correlation for SCLC Cohort3. The evaluation of the rest of the 8 antibodies didn't result in any significant correlation with survival. For different cohorts any of the autologous antibody responses doesn't show correlation with survival due to tumor heterogeneity.

We have also investigated correlation of seropositivity and survival by Kaplan-Meier method. SOX2 antibody seropositivity according to the M+2SD cut-off correlate with overall survival for SCLC Cohort2 whereas there is no significant correlation for SCLC Cohort3. So that this confirms our other analysis results.

We analyzed the correlation between seropositivity and antigen expression in the tumor tissue for SOX2. We have identified a correlation between seropositivity against SOX2 and SOX2 protein expression intensity in the SCLC tissues. However SOX2 seropositivity was not correlating with SOX2 antigen expression frequency. Our ELISA data was also confirmed with Western blotting.

Our optimized protocols were utilized for identification of autologous antibodies for NBD; an autoimmune disease. Mtch1 autologous antibodies were identified by screening of PA with NBD patients' sera in parallel with healthy controls and the discriminative power of Mtch1 was confirmed with ELISA. Mtch1 is an apoptosis related protein and it is highly probable that the autologous antibodies are relevant in terms of pathogenicity of NBD.

We have also screened PAs with BD sera and compared the results with healthy control screening results. We have identified 8 proteins as the most valuable antigens capable of discriminating BD from the healthy controls. The most valuable 8 clones were analyzed by MC for the SAS values for identification of the best panels. The most valuable panel was formed by 4 autologous antibodies and reached a sensitivity of 42% at 100% specificity.

Chapter 7

7 DISCUSSION AND CONCLUSION

7.1 Method Development and Protein Arrays

The method we have developed and optimized is very advantageous for identification of cancer specific autologous antibodies by comparison of cancer and healthy control screened PAs since both signal intensity stability and reproducibility was achieved by utilizing Q-Dots for signal generation. PA developer has recommended chemiluminescent or organic fluorophores based signal based detection methods for screening PAs. However chemiluminescent signals are dynamic and have a short duration thus not suitable for parallel PA screening to identify cancer specific autologous antibodies [71].

Q-Dots were repeatedly reported to have 100-1000 times more chemical stability and 10-20 times more brightness/unit compared to organic fluorophores [72, 73]. Organic fluorophores have a narrow absorption spectrum which results in a narrow range of emission. Also the organic dyes do not have a sharp symmetric emission peak which is further broadened by a red-tail [74, 75]. In contrast, Q-dots have a broader excitation spectra and a narrow more sharply defined emission peak [74]. Q-dots also have a brighter emission and a higher signal to noise ratio compared with organic dyes [76]. The brightness of Q-dots compared to organic dyes are 10-20 times brighter [77]. Q-dots are stable fluorophores due to their inorganic composition which reduces the effect of photobleaching compared to organic dyes [74, 75]. The fluorescence time for Q-dots is about 10 to 40 ns which is longer than the fluorescence of a few nanoseconds of organic dyes [74, 75, 78]. The inorganic composition of quantum dots also make them more robust toward degradation which contributes to their longevity [79]. The large Stokes shift (difference between peak absorption and peak emission wavelengths) reduces autofluorescence which increases sensitivity [79]. Thus Q-Dots provided comparability of both cancer screening and healthy control screening results. The

optimization efforts for each step of the screening resulted in maximum signal/background ratio assessment.

There are a couple of protein arrays suitable for screening for autologous anti-tumor antibody identification on the market. These are ProtoArray from ThermoFisher Scientific including more than 9000 proteins [80-82], Immunome protein arrays from Sengenics including 1636 proteins [83, 84] and RayBiotech protein arrays including 487 proteins [85]. There is also a recent protein array from Cambridge Protein Arrays including 19000 proteins [86], however this protein array was not in the market when we planned performing PA screening. So Imagenes protein arrays included more than 27000 clones which had the highest protein diversity content when we decided to perform PA screening. Also fetal brain and testis are the 2 tissues that have the highest amount of mRNA diversity. Besides SCLC has a neuroectodermal origin and screening FB PA increased the possibility of identifying the maximum possible number of autologous antibodies.

7.2 PA Screening Evaluations

According to the PA screening results the signals selected from SCLC screenings were stronger than the signals selected from Healthy screenings for all of Testis and Fetal Brain PAs. This indicates that the strong signals selected from SCLC screening were cancer specific since SCLC is an immunogenic cancer and also constitutes a quality control for the PA screenings since we aimed to identify SCLC specific signals.

PA screening results' evaluations revealed some outputs which constitute quality control characteristics for the procedure:

- 1) We have identified proteins which are previously known to be eliciting autologous anti-tumor antibodies, which are SOX2, HuC and HuD.
- 2) We have identified clones from the same screening containing different cDNA forms of the same gene, which are CCDC9 and YBX1 for Testis PA screening and SOX2, HuC, CHMP5 and EEF1A1 from Fetal Brain PA screening.
- 3) CIC was identified from both Testis and Fetal Brain PA screenings.

7.3 Contradictions between PA and custom PA Screening Evaluations

7.3.1 Clones not recognized by pool but recognized by individual sera

According to the custom PA screening evaluations not only the clones selected from the SCLC pool screening but also from Ovarian, Gastric and Colorectal pool screenings were among the most valuable custom PA clones for diagnosing SCLC. This can be explained by unpredictable reactivities of autologous antibodies with other antibodies of other serum in the pool. The autoantibodies reacting with the clones selected from non-SCLC pool screenings were probably precipitated by other antibodies from other sera, recognizing these autoantibodies as non-self in the pool. So these clones were not recognized with the pool but are recognized with individual sera.

7.3.2 Clones recognized by pool but not recognized by individual sera

Pooling also explains why only 10 of all 40 clones selected from SCLC pool screening were recognized by autologous antibodies of individual sera during screening of custom PAs. This can be explained by formation of pool specific structures between antibodies from different sera and recognition of some clones by these structures which causes formation of false positive signals.

7.4 Novel autologous antibody identification and their diagnostic value

For most of the cancers we have screened with PA (except Colorectal for FB and Ovarian for Testis), the number of cancer specific positive clones was always bigger than the common cancer positive clones. This indicates that cancer specific autologous antibodies are valuable biomarkers to be utilized for specific cancer diagnosis and early diagnosis.

Panel formation and evaluation of the panel for diagnosis increases sensitivity of the method. So that inclusion of valuable antigens in the panel can increase the diagnostic power up to a certain level. However our data suggests that after a certain point sensitivity cannot be increased further most probably due to a general immune-suppression in the cancer patients. We have observed around 40% immune-suppression rate for all cancer types we have screened that patients have no antibody response against any of the antigens on custom PAs.

Some of the clones in custom PA coincidentally behaved like a positive control since around 70% of the healthy controls were seropositive to these clones (data not shown). For cancer patients this percentage is around 45% which indicated an immune-suppression. So that autologous antibodies have a diagnostic limitation as biomarkers. Immune-suppression is a well known mechanism utilized by the tumor for coping with the immune system during its progress [87].

For diagnosing the cancer patients in the immune-suppressed group alternative biomarkers can be utilized like protein, miRNA, expression and mutation/SNPs. Complementation with any of these biomarkers can result in increase of panel sensitivity for the biomarker panel to be utilized in clinics.

Currently the only kit available on the market for lung cancer diagnosis is Oncimmune EarlyCDT-Lung. The kit has a sensitivity of 40% at a specificity of 90% for lung cancer [26]. These values becomes 55% and 90% for SCLC respectively. Over 98% of the nodules identified by screening with low-dose CT are actually false positives due to its relatively low specificity. Combined application of the test with low-dose CT has the potential for decreasing the false positivity rate of low-dose CT. Recently application of additional metrics based on the antibody titration curves were shown to increase specificity of the test.

We have reached a sensitivity value of 55% at 90% specificity with the antigen panel of SOX2, p53, POLB and C11orf20 by our ELISA data which is identical to EarlyCDT-Lung's diagnostic power for SCLC. However the PC method we have developed through this thesis work exceeds the diagnostic power of EarlyCDT-Lung only by combined evaluation of SOX2 and p53. So that PC method harbors good potential to exceed the diagnostic potential of low-dose CT and further studies needs to be performed to evaluate this potential.

7.5 The literature survey for the most valuable 6 antigens

A comprehensive Pubmed search was performed to analyze the 6 autologous antibodies for assessing the knowledge already present for their relevance in tumor immunology and it is summarized below.

7.5.1 SOX2

Anti-SOX2 autoantibodies were first identified in SCLC by SEREX at high titer by our group; even detectable in 1:10⁶ serum dilution. [25] Major B-cell epitopes of SOX2 were identified to be recognized by prostate, glioblastoma and lung cancer serum samples in a separate study [88]. SOX2 autoantibodies were also shown to be more effective than assays of serum tissue polypeptide-specific antigen, carcinoembryonic antigen, CA 125, and CA 15-3 in discriminating malignant disease from benign for breast cancer [89]. Anti-SOX2 autologous antibodies were also shown to be elicited in 7.7% of multiple myeloma patients [90]. So that SOX2 autologous antibodies seems like cancer specific, not SCLC specific.

7.5.2 p53

p53 autologous antibodies were identified in multiple cancers. A meta-analysis was performed for evaluation of autologous antibodies identified till present time for esophageal cancer revealed the most frequently assessed autoantibody (in 45 different research articles) was anti-p53 among 35 different autologous antibodies, which was tested in 17 studies and for 15 studies of which a meta-analysis was conducted to comprehensively evaluate the diagnostic value [91]. In addition to the antitumor functions of p53, accumulating evidence using systemic p53-deficient mice suggests that p53 suppresses autoimmunity. It was shown that aged p53-knock-out mice spontaneously developed inflammatory lesions in various organs, including lung, liver, stomach, thyroid gland, submandibular gland, and kidney. Additionally, anti-nuclear Abs and autoantibodies against gastric parietal cells were detected in p53-knock-out mice but not in p53 wild-type mice. These findings suggest a role for a link between carcinogenesis and autoimmunity [92].

7.5.3 C11orf20

C11orf20 (TEX 40) is an uncharacterized gene though it is conserved in the mammalian lineage. There is only one paper in Pubmed showing a recurrent gene fusion of the first 2 exons of Estrogen Receptor Related Alpha (ESRRA) to the 3. , 4. , or 5. exon of C11orf20 both on chromosome and transcripts in 15% of serous ovarian cancer patients' tumors which might be responsible for the development of the cancer at least in a subset of these patients. Normally C11orf20 and ESRRA are consecutively positioned on Chromosome 11 [93]. We have screened 54 Ovarian Cancer patients' sera for anti-C11orf20 antibodies (data not shown), however we couldn't find any significant difference between the

seropositivity rates of these patients and healthy controls. This can be explained by a recent study performed with 197 serous ovarian cancer patients which couldn't validate the ESRRA/C11orf20 fusion in any of the patients thus indicated that it is a rare event for this cancer type [94]. So that C11orf20 is a novel autologous antibody identified for SCLC for the first time by this study.

7.5.4 POLB

Many studies address POLB relation with cancer since it is well characterized and the smallest DNA polymerase which is the main enzyme in the base excision repair pathway. Approximately 30% of human tumors sequenced to date have mutations in POLB gene which are absent in the surrounding healthy tissue. These mutations are also shown to be responsible for transforming cells when expressed; one example is the D160N identified in gastric cancer [95]. A significant number of studies were performed for correlating certain mutations and SNPs with cancer development. 567 single nucleotide polymorphisms (SNPs) have been identified so far, 34 of which are in the coding region and 22 are in the non-coding regions of the exons [96]. Only two of them are confirmed by larger cohorts and one of them; P242R, is defined to be causing loss of enzyme function [96]. Among 652 polymorphisms in 85 genes involved in maintenance of genetic stability two genes; one of which is POLB was shown to be significantly associated with bladder cancer in a study comparing 201 incident bladder cancer cases and 326 controls [97]. A similar study was also performed in 1150 cases of urinary bladder transitional cell carcinoma and 1149 controls and again an intronic SNP of POLB was found to be significantly associated with bladder cancer [98]. Another study identified an association between POLB SNPs with lung cancer [99]. POLB overexpression was observed in esophageal cancer tissues compared to the adjacent non-cancer tissue [100]. A mutation impairing the catalytic activity of POLB was shown to be associated with low differentiated cervical cancer tissues compared to moderate and high ones [101]. Among 977 non-synonymous SNPs (nsSNP) investigated in 425 chronic lymphocytic leukemia patients only 9 were significantly correlating both with worse overall and progression-free survival one of which is POLB P242R [102]. In a collection of 134 colorectal cancer patients' tumors POLB showed coding region mutations, which caused reduced enzyme activity, in 40% of the samples [103]. High POLB expression is shown to be correlating with poorer prognosis in colorectal cancer patients [104]. Among 109 nsSNPs in 50 DNA repair genes were investigated in 700 lung cancer patients and 15 of them were significantly correlated with worse overall survival one of which again POLB P242R [105]. A study performed with 378

pancreatic cancer cases identified a significant correlation between two POLB SNPs; A165G and T2133C and overall survival [106]. So that POLB autologous antibodies were not reported to date in cancer patients' sera. This is the first study reporting the POLB autologous antibodies in SCLC.

7.5.5 ALMS1

Alstrom syndrome 1 (ALMS1) is a novel gene which is ubiquitously expressed and has multiple splice variants, was reported to be mutated in Alstrom syndrome which is a rare genetic disease. Childhood obesity, insulin resistance, type 2 diabetes, retinal degeneration, cardiomyopathy, kidney and liver disease, sensorineural hearing loss and infertility are well-established features of Alstrom syndrome [107-110]. Extensive fibrosis is observed in patients with Alstrom syndrome have in almost all tissues causing organ failure which is the cause of death in most cases [111]. ALMS1 is widely expressed and located to centrosomes and to the base of cilia. The protein is implicated in ciliary function, cell cycle control and intracellular transport [107]. ALMS1 mutations are reported to cause impaired function rather than developmental anomalies which are thought to be causing the syndrome [110]. It was shown in mice that cochlear histopathologies are associated with the ciliopathy that ALMS1 is implicated in the planar cell polarity signaling and the loss of outer hair cells causes the majority of the hearing loss [108]. Based on studies performed with dermal fibroblasts from patients with Alstrom syndrome revealed that these fibroblasts display cytoskeletal abnormalities and migration impairment, up-regulate the expression and production of collagens and as opposed to increase in the cell cycle duration are more resistant to apoptosis [111] Male mice with mutated ALMS1 are infertile because of a progressive germ cell death followed by an almost entire block of development at spermatogenesis [109]. These data suggest that ALMS1 has critical roles in neuronal function and spermatogenesis so that its expression is critical in nervous and testis tissue which might be a link to explain why it is immunogenic in SCLC since SCLC has a neuroectodermal origin. Also this study is the first to show ALMS1 autologous antibodies in SCLC patients.

7.5.6 TUSC1

TUSC1 was first characterized by loss of heterozygosity analysis on chromosome 9p which was altered in numerous cancers and indicated existence of tumor suppressor genes on that location. In a study 30 NSCLC and 12 SCLC cell lines were screened for homozygous deletions and TUSC1 was identified. Downregulation of TUSC1 was observed in all lung cancer cell lines compared to tissues [112]. Higher expression levels of TUSC1 correlated with overall survival according to immunohistochemistry analysis of a primary lung cancer tissue microarray. TUSC1 overexpression was shown to result in growth inhibition both in vitro and in vivo [113]. TUSC1 might possibly be a tumor suppressor gene like p53 which is mutated in most tumors and p53 mutations correlate with autoantibody elucidation.

7.6 Prognostic analysis of autologous antibodies

We were able to identify a correlation between SOX2 and TUSC1 seropositivities and survival for SCLC Cohort2 and ALMS1 seropositivity and survival for SCLC Cohort3. The inconsistency in terms of survival and seropositivity correlation between different cohorts can be explained by tumor heterogeneity. However the correlation of SOX2 seropositivity and survival can be explained by the possibility that the immune response against SOX2 is eliminating the aggressive character of the tumor since SOX2 protein expression has been related to more aggressive tumors in several studies [114-117]. Also upregulation of this gene is known to enhance tumor cell proliferation [118]. SOX2 overexpression has also been shown to be essential for lung cancer stem cell function [119, 120]. So that it is possible that autologous antibodies are associated with good survival and the two contradictory effects cancel each other's effect. The extent of this elimination and the equilibrium probably forms the difference between cohorts in terms of seropositivity and survival correlation.

On the other hand downregulation of TUSC1 was observed in all lung cancer cell lines compared to tissues [112]. Higher expression levels of TUSC1 correlated with overall survival according to immunohistochemistry analysis of a primary lung cancer tissue microarray. TUSC1 overexpression was shown to result in growth inhibition both in vitro and in vivo [113]. TUSC1 might possibly be a tumor suppressor gene. So that immunity against TUSC1 probably develops against its impaired and overexpressed form which is associated with dysregulated tumor suppressor function thus eliminates aggressive tumor cell population and favors good prognosis.

ALMS1 autologous antibodies are associated with worse overall survival. Most probably ALMS1 ectopic expression is the cause of the antibody response in SCLC. Since ALMS1 has critical roles in neuronal function and spermatogenesis so that its expression is critical in nervous and testis tissue which might be a link to explain why it is immunogenic in SCLC since SCLC has a neuroectodermal origin [111]. ALMS1 expression might possibly indicate less differentiated tumor which can explain the autologous antibody presence and poor prognosis.

7.7 Correlates of antigen expression, presence of autologous antibodies and clinical parameters

The illumination of the mechanisms underlying antitumor immune responses is critical as such responses can be ultimately boosted if their beneficial affect can be proven. This is the first study where autologous anti-tumor antibody responses against SOX2 have been correlated with tumor antigen expression in SCLC. It is known that autologous antibodies can be elicited against either upregulated, mutated or foreign proteins. SOX2 is amplified and thus upregulated, but not mutated in SCLC [118]. In patients with monoclonal gammopathy of undetermined significance (MGUS), anti-SOX2 T cell responses were found to be directed against a very small percentage of tumor cells which were of the clonogenic type [68]. In the same study, patients with anti-SOX2 T cell showed significantly better overall survival. In contrast to the case with MGUS, in SCLC, SOX2 immune responses, as measured by the presence of antibody, have been observed to correlate with better outcome only in one of the three cohorts studied so far [69, 70, 121]. We believe there could be several explanations for this. Firstly, as demonstrated in the MGUS study, anti-SOX2 T-cell and antibody responses do not overlap completely; therefore, if T-cell responses determine outcome more so than antibodies, then this awaits to be demonstrated for SCLC. Secondly, no study to date has shown an association between worse outcome, or prognosis with SOX2 seropositivity. In contrast, SOX2 protein expression has been related to more aggressive tumors in several studies [114-117] and the upregulation of this gene is known to enhance tumor cell proliferation [118]. In addition, SOX2 overexpression has been shown to be essential for lung cancer stem cell function [119, 120]. It is possible that antibody responses are associated with improved outcome, but given the fact that SOX2 expression has the opposite effect, the two cancel out each other. Larger and comprehensive studies might further clarify these matters.

A very strong evidence that anti-tumor immune responses can effectively eliminate selected cells in humans comes from the observation that anti-tumor responses against antigens expressed in tumors but also in normal tissues, and in particular in Purkinje cells of the cerebellum, are able to completely eliminate these cells, without destroying surrounding tissues [65]. The amount of antigen expressed by a given cell has been shown to determine the type of T cell response directed against it [122]. HER-2/neu specific immunity, for example, has been shown to depend on the antigens' expression level [123]. Therefore, it is possible that only cells with intense SOX2 expression were able to induce immune responses against them in the cohort we studied, potentially resulting in their loss, similar to the case with Purkinje cells in patients with anti-HuD antibodies. The fact that in our study SOX2 seropositive patients' tumors on occasion contained very few cells could be reflecting the fact that such tumors undergo immune-editing [124], and reach a state of equilibrium between the tumor and the immune response following a loss of most SOX2 expressing cells [125]. Another explanation could be that those cells in tumors from patients who are seropositive represent clonogenic cells which elicit an immune response as observed for MGUS patients [68]. In glioma the intensity and not the frequency of SOX2 expression was shown to be an indicator of a more stem-like phenotype [126]. On the other hand, the weak but diffuse staining in a number of tumors we studied might reflect the presence of SOX2 expressing non-clonogenic cells. This is likely as we and others have found morphologically normal bronchial epithelium to be frequently positive for SOX2 [127], and is also supported by experiments showing that the transfection of a second gene in addition to SOX2 is required for tumorigenic transformation in some models [128, 129].

We report, for the first time, a relation between SOX2 protein expression characteristics and anti-SOX2 antibody responses in patients with SCLC. Although we find no correlation between outcome or clinical measures with frequency or intensity of SOX2 expression, we observe that SOX2 seropositivity associates with better prognosis. Tumors from patients with SOX2 antibody generally contain strongly staining cells, in contrast to tumors from seronegative patients. This suggests that the intensity of SOX2 expression might be critical in eliciting an anti-SOX2 immune response. The fact that several tumors have very low numbers of such cells suggests that SOX2 expressing cells could have been eliminated over the course of disease, which is in support of active tumor immune-surveillance and immune-editing in these patients.

7.8 Evaluation of autologous antibodies in autoimmune disease

A number of autoantibodies have been described in serum and/or CSF samples of NBD patients. Most of these are directed against stress related proteins, such as HSP-60, HSP-65, HSP-70, α B-crystallin and stress-induced-phosphoprotein 1. Antibodies to PINK1, α -enolase, cyclic citrullinated peptide and *S. cerevisiae* antigens have also been identified. However, all of these antibodies are found only in a small fraction (5–35%) of NBD patients and they can also be frequently detected in patients with other neuroinflammatory diseases lowering their values as a diagnostic [130-138]. Annexin-V antibodies are highly prevalent in NBD and BD patient cohorts [139], but can be detected in a plethora of rheumatological disorders [140] and the prevalence in other neuroinflammatory disorders is currently unknown. The novel Mtch1 antibody described in this study does not only add a new member to the list of NBD-associated antibodies but also appears to be a potential diagnostic biomarker with its high frequency in NBD patients and low prevalence in other neuroinflammatory diseases that are in the differential diagnosis list of NBD such as multiple sclerosis.

Presently, the only reliable diagnostic test for BD and NBD is the pathergy test, which also constitutes a part of the BD diagnostic criteria. However, problems with standardizing the induction method, needle size and type as well as the method of assessment of the response have limited the usefulness of this test in the clinical setting. Also, the prevalence of a positive pathergy test in BD varies among countries and test centers. Sensitivity of the pathergy test ranges between 30% and 45% at specificities of 87% to 100% [141-144]. The sensitivity and specificity values of Mtch1 antibodies for NBD and BD patients in our study were easily comparable with those of pathergy test. As a matter of fact, the frequency of Mtch1 antibody seropositivity in our NBD cohort (56.3%) was higher than that of pathergy test positivity (53.1%). Therefore, ELISA-based Mtch1 antibody measurement might potentially be utilized as a reliable, easy-to-use, noninvasive and standard diagnostic method.

NBD almost always develops several years after the onset of BD [145] and thus NBD patients are generally under immunosuppressive treatment. Since all of our NBD patients were under long-term follow-up as BD patients, we could not compare Mtch1 antibody levels in naïve and immunosuppressed NBD patients. Nevertheless, this comparison was made among BD patients. Although non-immunosuppressed BD patients showed trends towards exhibiting slightly higher Mtch1 antibody levels than immunosuppressed BD patients, this difference did not reach statistical significance, suggesting that measurement of Mtch1

antibody is not significantly affected by immunosuppression and lower Mtch1 antibody seropositivity rates in the BD group is not related with the differences in treatment status. However, for a better assessment of the specificity of Mtch1 antibody, patients with non-inflammatory CNS disorders as well as other vasculitic–rheumatological disorders need to be studied.

Based on the molecular mimicry between stress induced proteins and certain proteins expressed by microorganisms, it has recently been proposed that HSP antibodies develop as a result of the immune reaction against invading pathogens and coincidentally crossreact with human tissue thus causing BD symptoms [146]. However, our extensive search in National Center for Biotechnology Information GenBank using BLAST and CLC Main Work bench software has failed to find any considerable identity between Mtch1 and proteins of a wide range of microorganisms (data not shown). Moreover, Mtch1 antibody's association with NBD disease severity and apoptotic cell death rates suggests that Mtch1 antibody is not a bystander side effect of autoinflammation and might have certain pathogenic functions in NBD pathogenesis. It is well known that, in BD, an increased occurrence of apoptotic cell death is observed in parenchymal lesions of CNS as well as other involved tissues [147]. Notably, both Mtch1 and recently identified BD autoantigen annexin-V are associated with apoptosis and levels of antibodies to both antigens are correlated with disease severity [139, 140, 148, 149]. Mtch1 is a proapoptotic protein, Mtch1 antibodies tend to occur in patients with a more severe disease course and patients with Mtch1 antibodies exhibit reduced intensity of apoptotic cell death, altogether suggesting that Mtch1 antibodies might plausibly be developing as a protective mechanism to reduce and neutralize the tissue damage afflicted by BD associated autoinflammation.

Antibodies to progranulin, a protein associated with frontotemporal dementia, have recently been discovered in vasculitis patients [150]. Notably, Mtch1 is closely associated with presenilin, the dysfunction of which might cause Alzheimer's disease [148, 149]. Whether these recent findings point to a possible link between neurodegeneration and neuroinflammation requires to be scrutinized in future studies. In conclusion, Mtch1 antibody seropositivity appears to have a considerable sensitivity and specificity for NBD and BD. The association between Mtch1 antibody levels and clinical and apoptotic parameters of NBD suggests that Mtch1 antibody might have a pathogenic action. Therefore, further characterization of the functional role of Mtch1 antibody is warranted.

Chapter 8

8 FUTURE PERSPECTIVES

8.1 Application of PC to the PA data

During the PhD study we have applied the PC method only to two autologous antibodies which are SOX2 and p53. Only these 2 autologous antibodies resulted in 82% sensitivity at 90% specificity which is comparable to the diagnostic power of low-dose CT. We are planning to apply the method to all the valuable 23 clones we have identified as a result of the PA and custom PA screenings. Since we have shown that MC results are consistent with SVM, we will perform MC to the values obtained by PC. This will result in identification of the best antigen panel and the ideal cut-offs. As a result an ROC will be plotted and the AUC will be calculated. For validation of the results we will replicate the PA with the selected clones in the lab. We have already obtained the protocol for PA production from the references shared by the manufacturer. After replication of the PA we will validate the PC results first for a representative group of sera. We will perform several optimization steps for fine-tuning. After the optimizations are complete we will screen all the 3 SCLC and 2 healthy control cohorts with the PA. We will obtain the values after application of PC and evaluate these values by MC. As a result we will determine the ideal cut-offs for each antibody and plot ROC for each cohort. We will calculate AUC values and compare the cohorts. We expect to obtain better sensitivity and specificity values compared to chest X-ray and low-dose CT.

8.2 Diagnostic kit development

If we obtain better sensitivity and specificity values compared to chest X-ray and low-dose CT we will apply for grants for development of a kit based on the autologous antibodies in the final panel and PC method. We will perform a SWOT analysis and elaborate a business plan based on business canvas model. We have already established a start-up (PRZ BioTech) for this purpose and the kit will be developed to be a product of the start-up.

8.3 Colorectal, Gastric and Ovarian cancers

We have already performed PA screenings with pooled serum for colorectal, gastric and ovarian cancers besides SCLC. After we complete the study with PC for SCLC and replicating the PA in-house and all the analysis for obtaining a powerful diagnostic panel we will focus on these other cancer types and utilize the same strategy for these cancers to identify a valuable antibody panel for diagnosis.

8.4 Utilization of antibodies for early diagnosis of cancer

After we identify the best antibody panel based on the PC values we will also analyze the biomarkers for benefit in early diagnosis. Seperate grant applications will be performed for this purpose. In context of such studies we will collaborate with lung cancer clinicians and collect samples from individuals who have high-risk for developing lung cancer. The high-risk definition will be people who are 1) older than 50 and 2) actively smoking tobacco products. The serum samples will be collected from these individuals periodically; in every 3 months. These sera will be screened for the most valuable antibodies combined in the panel in comparison to the healthy controls and SCLC patients. The results will be evaluated according to the already defined MC cut-offs. The benefit of the autologous antibodies for early diagnosis will be assessed and a seperate kit development strategy will be elaborated.

9 BIBLIOGRAPHY

1. Society, A.C., *Global Cancer Facts & Figures* r. Edition, Editor. 2015: Atlanta.
2. Bach, P.B., et al., *Variations in lung cancer risk among smokers*. J Natl Cancer Inst, 2003. **95**(6): p. 470-8.
3. Boyle, P., B.E. Levin, and *World Cancer Report*. 2008, IARC: Lyon, France.
4. Field, J.K. and S.W. Duffy, *Lung cancer screening: the way forward*. Br J Cancer, 2008. **99**(4): p. 557-62.
5. Felip, E., et al., *ESMO Minimum Clinical Recommendations for diagnosis, treatment and follow-up of small-cell lung cancer (SCLC)*. Ann Oncol, 2005. **16 Suppl 1**: p. i30-1.
6. Lauer, M.S., et al., *[18F]Fluorodeoxyglucose uptake by positron emission tomography for diagnosis of suspected lung cancer: impact of verification bias*. Arch Intern Med, 2007. **167**(2): p. 161-5.
7. Jemal, A., et al., *Global patterns of cancer incidence and mortality rates and trends*. Cancer Epidemiol Biomarkers Prev, 2010. **19**(8): p. 1893-907.
8. Koletsis, E.N., et al., *Current role of surgery in small cell lung carcinoma*. J Cardiothorac Surg, 2009. **4**: p. 30.
9. Lim, E., et al., *The role of surgery in the treatment of limited disease small cell lung cancer: time to reevaluate*. J Thorac Oncol, 2008. **3**(11): p. 1267-71.
10. Brock, M.V., et al., *Surgical resection of limited disease small cell lung cancer in the new era of platinum chemotherapy: Its time has come*. J Thorac Cardiovasc Surg, 2005. **129**(1): p. 64-72.
11. Annema, J.T., et al., *Endoscopic ultrasound-guided fine-needle aspiration in the diagnosis and staging of lung cancer and its impact on surgical staging*. J Clin Oncol, 2005. **23**(33): p. 8357-61.
12. Delgado, P.I., M. Jorda, and P. Ganjei-Azar, *Small cell carcinoma versus other lung malignancies: diagnosis by fine-needle aspiration cytology*. Cancer, 2000. **90**(5): p. 279-85.
13. Chin, R., Jr., et al., *Increasing use of bronchoscopic needle aspiration to diagnose small cell lung cancer*. Mayo Clin Proc, 2000. **75**(8): p. 796-801.
14. Tondini, M. and A. Rizzi, *Small-cell lung cancer: importance of fiberoptic bronchoscopy in the evaluation of complete remission*. Tumori, 1989. **75**(3): p. 266-8.
15. Cristallini, E.G., et al., *Fine needle aspiration biopsy in the diagnosis of intrathoracic masses*. Acta Cytol, 1992. **36**(3): p. 416-22.
16. Hassanein, M., et al., *The state of molecular biomarkers for the early detection of lung cancer*. Cancer Prev Res (Phila), 2012. **5**(8): p. 992-1006.
17. Heo, C.K., Y.Y. Bahk, and E.W. Cho, *Tumor-associated autoantibodies as diagnostic and prognostic biomarkers*. BMB Rep, 2012. **45**(12): p. 677-85.
18. Murphy, M.A., J.J. O'Leary, and D.J. Cahill, *Assessment of the humoral immune response to cancer*. J Proteomics, 2012. **75**(15): p. 4573-9.
19. DeLeo, A.B., et al., *Cell surface antigens of chemically induced sarcomas of the mouse. I. Murine leukemia virus-related antigens and alloantigens on cultured fibroblasts and sarcoma cells: description of a unique antigen on BALB/c Meth A sarcoma*. J Exp Med, 1977. **146**(3): p. 720-34.
20. DeLeo, A.B., et al., *Detection of a transformation-related antigen in chemically induced sarcomas and other transformed cells of the mouse*. Proc Natl Acad Sci U S A, 1979. **76**(5): p. 2420-4.

21. Uyttenhove, C., J. Van Snick, and T. Boon, *Immunogenic variants obtained by mutagenesis of mouse mastocytoma P815. I. Rejection by syngeneic mice*. J Exp Med, 1980. **152**(5): p. 1175-83.
22. Svane, I.M., et al., *Chemically induced sarcomas from nude mice are more immunogenic than similar sarcomas from congenic normal mice*. Eur J Immunol, 1996. **26**(8): p. 1844-50.
23. Traversari, C., et al., *A nonapeptide encoded by human gene MAGE-1 is recognized on HLA-A1 by cytolytic T lymphocytes directed against tumor antigen MZ2-E*. J Exp Med, 1992. **176**(5): p. 1453-7.
24. Sahin, U., O. Tureci, and M. Pfreundschuh, *Serological identification of human tumor antigens*. Curr Opin Immunol, 1997. **9**(5): p. 709-16.
25. Gure, A.O., et al., *Serological identification of embryonic neural proteins as highly immunogenic tumor antigens in small cell lung cancer*. Proc Natl Acad Sci U S A, 2000. **97**(8): p. 4198-203.
26. Murray, A., et al., *Technical validation of an autoantibody test for lung cancer*. Ann Oncol, 2010. **21**(8): p. 1687-93.
27. Chapman, C.J., et al., *Autoantibodies in lung cancer: possibilities for early detection and subsequent cure*. Thorax, 2008. **63**(3): p. 228-33.
28. Chapman, C., et al., *Autoantibodies in breast cancer: their use as an aid to early diagnosis*. Ann Oncol, 2007. **18**(5): p. 868-73.
29. Zhong, L., et al., *Profiling tumor-associated antibodies for early detection of non-small cell lung cancer*. J Thorac Oncol, 2006. **1**(6): p. 513-9.
30. Zhang, J.Y., *Tumor-associated antigen arrays to enhance antibody detection for cancer diagnosis*. Cancer Detect Prev, 2004. **28**(2): p. 114-8.
31. Zhang, J.Y., et al., *Enhancement of antibody detection in cancer using panel of recombinant tumor-associated antigens*. Cancer Epidemiol Biomarkers Prev, 2003. **12**(2): p. 136-43.
32. Tan, E.M., *Autoantibodies as reporters identifying aberrant cellular mechanisms in tumorigenesis*. J Clin Invest, 2001. **108**(10): p. 1411-5.
33. Robertson, J.F.R., et al., *Autoantibodies in early breast cancer*. Journal of Clinical Oncology, 2005. **23**(16): p. 16s-16s.
34. Li, Y., et al., *p53 autoantibodies predict subsequent development of cancer*. Int J Cancer, 2005. **114**(1): p. 157-60.
35. Trivers, G.E., et al., *Anti-p53 antibodies in sera from patients with chronic obstructive pulmonary disease can predate a diagnosis of cancer*. Clin Cancer Res, 1996. **2**(10): p. 1767-75.
36. Ramos-Yeo, Y.L. and C.V. Reyes, *Myasthenic syndrome (Eaton-Lambert syndrome) associated with pulmonary adenocarcinoma*. J Surg Oncol, 1987. **34**(4): p. 239-42.
37. Darnell, R.B. and J.B. Posner, *Paraneoplastic syndromes involving the nervous system*. N Engl J Med, 2003. **349**(16): p. 1543-54.
38. Posner, J.B. and J. Dalmau, *Paraneoplastic syndromes*. Curr Opin Immunol, 1997. **9**(5): p. 723-9.
39. Maddison, P. and B. Lang, *Paraneoplastic neurological autoimmunity and survival in small-cell lung cancer*. J Neuroimmunol, 2008. **201-202**: p. 159-62.
40. Motomura, M., et al., *Incidence of serum anti-P/O-type and anti-N-type calcium channel autoantibodies in the Lambert-Eaton myasthenic syndrome*. J Neurol Sci, 1997. **147**(1): p. 35-42.
41. Monstad, S.E., et al., *Hu and voltage-gated calcium channel (VGCC) antibodies related to the prognosis of small-cell lung cancer*. J Clin Oncol, 2004. **22**(5): p. 795-800.
42. Bataller, L., et al., *Antibodies to Zic4 in paraneoplastic neurologic disorders and small-cell lung cancer*. Neurology, 2004. **62**(5): p. 778-82.
43. Bazhin, A.V., et al., *Low titre autoantibodies against recoverin in sera of patients with small cell lung cancer but without a loss of vision*. Lung Cancer, 2001. **34**(1): p. 99-104.

44. Lennon, V.A., et al., *Calcium-channel antibodies in the Lambert-Eaton syndrome and other paraneoplastic syndromes*. N Engl J Med, 1995. **332**(22): p. 1467-74.
45. Galanis, E., et al., *Neuronal autoantibody titers in the course of small-cell lung carcinoma and platinum-associated neuropathy*. Cancer Immunol Immunother, 1999. **48**(2-3): p. 85-90.
46. Saiz, A., et al., *Anti-amphiphysin I antibodies in patients with paraneoplastic neurological disorders associated with small cell lung carcinoma*. J Neurol Neurosurg Psychiatry, 1999. **66**(2): p. 214-7.
47. Chan, J.W., *Paraneoplastic retinopathies and optic neuropathies*. Surv Ophthalmol, 2003. **48**(1): p. 12-38.
48. Bataller, L. and J. Dalmau, *[Antineural antibodies and paraneoplastic neurologic syndromes]*. Neurologia, 2002. **17**(2): p. 85-96.
49. Graus, F., et al., *Anti-Hu-associated paraneoplastic encephalomyelitis: analysis of 200 patients*. Brain, 2001. **124**(Pt 6): p. 1138-48.
50. Adamus, G., G. Ren, and R.G. Weleber, *Autoantibodies against retinal proteins in paraneoplastic and autoimmune retinopathy*. BMC Ophthalmol, 2004. **4**: p. 5.
51. Shams'ili, S., et al., *Paraneoplastic cerebellar degeneration associated with antineuronal antibodies: analysis of 50 patients*. Brain, 2003. **126**(Pt 6): p. 1409-18.
52. Lubin, R., et al., *Serum p53 antibodies as early markers of lung cancer*. Nat Med, 1995. **1**(7): p. 701-2.
53. Trivers, G.E., et al., *Anti-p53 antibodies in sera of workers occupationally exposed to vinyl chloride*. J Natl Cancer Inst, 1995. **87**(18): p. 1400-7.
54. Cawley, H.M., et al., *Anti-p53 antibodies in patients with Barrett's esophagus or esophageal carcinoma can predate cancer diagnosis*. Gastroenterology, 1998. **115**(1): p. 19-27.
55. Chapman, C.J., et al., *Immunobiomarkers in small cell lung cancer: potential early cancer signals*. Clin Cancer Res, 2011. **17**(6): p. 1474-80.
56. Chapman, C.J., et al., *EarlyCDT(R)-Lung test: improved clinical utility through additional autoantibody assays*. Tumour Biol, 2012. **33**(5): p. 1319-26.
57. Chen, Y., et al., *Autoantibodies to Ca²⁺ binding protein Calnuc is a potential marker in colon cancer detection*. Int J Oncol, 2007. **30**(5): p. 1137-44.
58. Stockert, E., et al., *A survey of the humoral immune response of cancer patients to a panel of human tumor antigens*. J Exp Med, 1998. **187**(8): p. 1349-54.
59. Tan, H.T., et al., *Serum autoantibodies as biomarkers for early cancer detection*. FEBS J, 2009. **276**(23): p. 6880-904.
60. Ludwig, N., et al., *Is there a general autoantibody signature for cancer?* Eur J Cancer, 2012. **48**(15): p. 2451-61.
61. Old, L.J. and Y.T. Chen, *New paths in human cancer serology*. J Exp Med, 1998. **187**(8): p. 1163-7.
62. Chen, Y.T., et al., *A testicular antigen aberrantly expressed in human cancers detected by autologous antibody screening*. Proc Natl Acad Sci U S A, 1997. **94**(5): p. 1914-8.
63. Winter, S.F., et al., *Antibodies against autologous tumor cell proteins in patients with small-cell lung cancer: association with improved survival*. J Natl Cancer Inst, 1993. **85**(24): p. 2012-8.
64. Verschuuren, J.J., et al., *Paraneoplastic anti-Hu serum: studies on human tumor cell lines*. J Neuroimmunol, 1997. **79**(2): p. 202-10.
65. Peterson, K., et al., *Paraneoplastic cerebellar degeneration. I. A clinical analysis of 55 anti-Yo antibody-positive patients*. Neurology, 1992. **42**(10): p. 1931-7.
66. Mattioni, M., et al., *Prognostic role of serum p53 antibodies in lung cancer*. BMC Cancer, 2015. **15**: p. 148.
67. Sabater, L., et al., *SOX1 antibodies are markers of paraneoplastic Lambert-Eaton myasthenic syndrome*. Neurology, 2008. **70**(12): p. 924-8.
68. Spisek, R., et al., *Frequent and specific immunity to the embryonal stem cell-associated antigen SOX2 in patients with monoclonal gammopathy*. J Exp Med, 2007. **204**(4): p. 831-40.

69. Vural, B., et al., *Frequency of SOX Group B (SOX1, 2, 3) and ZIC2 antibodies in Turkish patients with small cell lung carcinoma and their correlation with clinical parameters*. Cancer, 2005. **103**(12): p. 2575-83.
70. Titulaer, M.J., et al., *SOX antibodies in small-cell lung cancer and Lambert-Eaton myasthenic syndrome: frequency and relation with survival*. J Clin Oncol, 2009. **27**(26): p. 4260-7.
71. Weeks, I., et al., *Acridinium esters as high-specific-activity labels in immunoassay*. Clin Chem, 1983. **29**(8): p. 1474-9.
72. Jaiswal, J.K., et al., *Long-term multiple color imaging of live cells using quantum dot bioconjugates*. Nat Biotechnol, 2003. **21**(1): p. 47-51.
73. Medintz, I.L., et al., *Quantum dot bioconjugates for imaging, labelling and sensing*. Nat Mater, 2005. **4**(6): p. 435-46.
74. Alivisatos, A.P., W. Gu, and C. Larabell, *Quantum dots as cellular probes*. Annu Rev Biomed Eng, 2005. **7**: p. 55-76.
75. Chan, W.C., et al., *Luminescent quantum dots for multiplexed biological detection and imaging*. Curr Opin Biotechnol, 2002. **13**(1): p. 40-6.
76. Kawasaki, E.S. and A. Player, *Nanotechnology, nanomedicine, and the development of new, effective therapies for cancer*. Nanomedicine, 2005. **1**(2): p. 101-9.
77. Gao, X., et al., *In vivo molecular and cellular imaging with quantum dots*. Curr Opin Biotechnol, 2005. **16**(1): p. 63-72.
78. Michalet, X., et al., *Quantum dots for live cells, in vivo imaging, and diagnostics*. Science, 2005. **307**(5709): p. 538-44.
79. Fu, A., et al., *Semiconductor nanocrystals for biological imaging*. Curr Opin Neurobiol, 2005. **15**(5): p. 568-75.
80. Gunawardana, C.G., N. Memari, and E.P. Diamandis, *Identifying novel autoantibody signatures in ovarian cancer using high-density protein microarrays*. Clin Biochem, 2009. **42**(4-5): p. 426-9.
81. Auger, I., et al., *New autoantigens in rheumatoid arthritis (RA): screening 8268 protein arrays with sera from patients with RA*. Ann Rheum Dis, 2009. **68**(4): p. 591-4.
82. Li, L., et al., *Identifying compartment-specific non-HLA targets after renal transplantation by integrating transcriptome and "antibodyome" measures*. Proc Natl Acad Sci U S A, 2009. **106**(11): p. 4148-53.
83. Gnjjatic, S., et al., *Seromic analysis of antibody responses in non-small cell lung cancer patients and healthy donors using conformational protein arrays*. J Immunol Methods, 2009. **341**(1-2): p. 50-8.
84. Adamus, G., *Autoantibody targets and their cancer relationship in the pathogenicity of paraneoplastic retinopathy*. Autoimmun Rev, 2009. **8**(5): p. 410-4.
85. Chen, G., et al., *Autoantibody profiles reveal ubiquitin 1 as a humoral immune response target in lung adenocarcinoma*. Cancer Res, 2007. **67**(7): p. 3461-7.
86. Hu, C.J., et al., *Identification of new autoantigens for primary biliary cirrhosis using human proteome microarrays*. Mol Cell Proteomics, 2012. **11**(9): p. 669-80.
87. Wang, Y., et al., *Hepatocellular Carcinoma Cells Induce Regulatory T Cells and Lead to Poor Prognosis via Production of Transforming Growth Factor-beta1*. Cell Physiol Biochem, 2016. **38**(1): p. 306-318.
88. Shih, J., et al., *Dominant B-cell epitopes from cancer/stem cell antigen SOX2 recognized by serum samples from cancer patients*. Am J Clin Exp Immunol, 2014. **3**(2): p. 84-90.
89. Sun, Y., et al., *SOX2 autoantibodies as noninvasive serum biomarker for breast carcinoma*. Cancer Epidemiol Biomarkers Prev, 2012. **21**(11): p. 2043-7.
90. Kobold, S., et al., *Patients with multiple myeloma develop SOX2-specific autoantibodies after allogeneic stem cell transplantation*. Clin Dev Immunol, 2011. **2011**: p. 302145.
91. Zhang, H., et al., *Serum autoantibodies in the early detection of esophageal cancer: a systematic review*. Tumour Biol, 2015. **36**(1): p. 95-109.

92. Kawashima, H., et al., *Tumor suppressor p53 inhibits systemic autoimmune diseases by inducing regulatory T cells*. J Immunol, 2013. **191**(7): p. 3614-23.
93. Salzman, J., et al., *ESRRA-C11orf20 is a recurrent gene fusion in serous ovarian carcinoma*. PLoS Biol, 2011. **9**(9): p. e1001156.
94. Micci, F., et al., *Low frequency of ESRRA-C11orf20 fusion gene in ovarian carcinomas*. PLoS Biol, 2014. **12**(2): p. e1001784.
95. Donigan, K.A., et al., *The human gastric cancer-associated DNA polymerase beta variant D160N is a mutator that induces cellular transformation*. DNA Repair (Amst), 2012. **11**(4): p. 381-90.
96. Sobol, R.W., *Genome instability caused by a germline mutation in the human DNA repair gene POLB*. PLoS Genet, 2012. **8**(11): p. e1003086.
97. Michiels, S., et al., *Genetic polymorphisms in 85 DNA repair genes and bladder cancer risk*. Carcinogenesis, 2009. **30**(5): p. 763-8.
98. Figueroa, J.D., et al., *Genetic variation in the base excision repair pathway and bladder cancer risk*. Hum Genet, 2007. **121**(2): p. 233-42.
99. Kzama, R., et al., *Lung cancer and DNA repair genes: multilevel association analysis from the International Lung Cancer Consortium*. Carcinogenesis, 2012. **33**(5): p. 1059-64.
100. Dong, Z.M., et al., *Difference in expression level and localization of DNA polymerase beta among human esophageal cancer focus, adjacent and corresponding normal tissues*. Dis Esophagus, 2006. **19**(3): p. 172-6.
101. Han, L.P., et al., *[Study on DNA polymerase beta gene mutation in human cervical cancer]*. Zhonghua Fu Chan Ke Za Zhi, 2003. **38**(10): p. 618-20.
102. Sellick, G.S., et al., *Scan of 977 nonsynonymous SNPs in CLL4 trial patients for the identification of genetic variants influencing prognosis*. Blood, 2008. **111**(3): p. 1625-33.
103. Donigan, K.A., et al., *Human POLB gene is mutated in high percentage of colorectal tumors*. J Biol Chem, 2012. **287**(28): p. 23830-9.
104. Iwatsuki, M., et al., *A platinum agent resistance gene, POLB, is a prognostic indicator in colorectal cancer*. J Surg Oncol, 2009. **100**(3): p. 261-6.
105. Matakidou, A., et al., *Genetic variation in the DNA repair genes is predictive of outcome in lung cancer*. Hum Mol Genet, 2007. **16**(19): p. 2333-40.
106. Li, D., et al., *Effects of base excision repair gene polymorphisms on pancreatic cancer survival*. Int J Cancer, 2007. **120**(8): p. 1748-54.
107. Collin, G.B., et al., *The Alstrom syndrome protein, ALMS1, interacts with alpha-actinin and components of the endosome recycling pathway*. PLoS One, 2012. **7**(5): p. e37925.
108. Jagger, D., et al., *Alstrom Syndrome protein ALMS1 localizes to basal bodies of cochlear hair cells and regulates cilium-dependent planar cell polarity*. Hum Mol Genet, 2011. **20**(3): p. 466-81.
109. Arsov, T., et al., *Fat aussie--a new Alstrom syndrome mouse showing a critical role for ALMS1 in obesity, diabetes, and spermatogenesis*. Mol Endocrinol, 2006. **20**(7): p. 1610-22.
110. Hearn, T., et al., *Subcellular localization of ALMS1 supports involvement of centrosome and basal body dysfunction in the pathogenesis of obesity, insulin resistance, and type 2 diabetes*. Diabetes, 2005. **54**(5): p. 1581-7.
111. Zulato, E., et al., *ALMS1-deficient fibroblasts over-express extra-cellular matrix components, display cell cycle delay and are resistant to apoptosis*. PLoS One, 2011. **6**(4): p. e19081.
112. Shan, Z., T. Parker, and J.S. Wiest, *Identifying novel homozygous deletions by microsatellite analysis and characterization of tumor suppressor candidate 1 gene, TUSC1, on chromosome 9p in human lung cancer*. Oncogene, 2004. **23**(39): p. 6612-20.
113. Shan, Z., et al., *TUSC1, a putative tumor suppressor gene, reduces tumor cell growth in vitro and tumor growth in vivo*. PLoS One, 2013. **8**(6): p. e66114.
114. Saigusa, S., et al., *Correlation of CD133, OCT4, and SOX2 in rectal cancer and their association with distant recurrence after chemoradiotherapy*. Ann Surg Oncol, 2009. **16**(12): p. 3488-98.

115. Wang, Q., et al., *Oct3/4 and Sox2 are significantly associated with an unfavorable clinical outcome in human esophageal squamous cell carcinoma*. Anticancer Res, 2009. **29**(4): p. 1233-41.
116. Sholl, L.M., et al., *Sox2 protein expression is an independent poor prognostic indicator in stage I lung adenocarcinoma*. Am J Surg Pathol, 2010. **34**(8): p. 1193-8.
117. Du, L., et al., *Sox2 nuclear expression is closely associated with poor prognosis in patients with histologically node-negative oral tongue squamous cell carcinoma*. Oral Oncol, 2011. **47**(8): p. 709-13.
118. Rudin, C.M., et al., *Comprehensive genomic analysis identifies SOX2 as a frequently amplified gene in small-cell lung cancer*. Nat Genet, 2012. **44**(10): p. 1111-6.
119. Xiang, R., et al., *Downregulation of transcription factor SOX2 in cancer stem cells suppresses growth and metastasis of lung cancer*. Br J Cancer, 2011. **104**(9): p. 1410-7.
120. Nakatsugawa, M., et al., *SOX2 is overexpressed in stem-like cells of human lung adenocarcinoma and augments the tumorigenicity*. Lab Invest, 2011. **91**(12): p. 1796-804.
121. Maddison, P., et al., *Autoimmunity to SOX2, clinical phenotype and survival in patients with small-cell lung cancer*. Lung Cancer, 2010. **70**(3): p. 335-9.
122. Daniels, M.A., et al., *Thymic selection threshold defined by compartmentalization of Ras/MAPK signalling*. Nature, 2006. **444**(7120): p. 724-9.
123. Goodell, V., et al., *Level of HER-2/neu protein expression in breast cancer may affect the development of endogenous HER-2/neu-specific immunity*. Mol Cancer Ther, 2008. **7**(3): p. 449-54.
124. Dunn, G.P., L.J. Old, and R.D. Schreiber, *The three Es of cancer immunoediting*. Annu Rev Immunol, 2004. **22**: p. 329-60.
125. Koebel, C.M., et al., *Adaptive immunity maintains occult cancer in an equilibrium state*. Nature, 2007. **450**(7171): p. 903-7.
126. Velpula, K.K., et al., *Cord blood stem cells revert glioma stem cell EMT by down regulating transcriptional activation of Sox2 and Twist1*. Oncotarget, 2011. **2**(12): p. 1028-42.
127. Chen, S., et al., *SOX2 gene regulates the transcriptional network of oncogenes and affects tumorigenesis of human lung cancer cells*. PLoS One, 2012. **7**(5): p. e36326.
128. Takahashi, K. and S. Yamanaka, *Induction of pluripotent stem cells from mouse embryonic and adult fibroblast cultures by defined factors*. Cell, 2006. **126**(4): p. 663-76.
129. Park, I.H., et al., *Reprogramming of human somatic cells to pluripotency with defined factors*. Nature, 2008. **451**(7175): p. 141-6.
130. Vural, B., et al., *Anti-neuronal and stress-induced-phosphoprotein 1 antibodies in neuro-Behcet's disease*. J Neuroimmunol, 2011. **239**(1-2): p. 91-7.
131. Vural, B., et al., *Seroreactivity against PTEN-induced putative kinase 1 (PINK1) in Turkish patients with Behcet's disease*. Clin Exp Rheumatol, 2009. **27**(2 Suppl 53): p. S67-72.
132. Lee, J.H., et al., *Human anti-alpha-enolase antibody in sera from patients with Behcet's disease and rheumatologic disorders*. Clin Exp Rheumatol, 2009. **27**(2 Suppl 53): p. S63-6.
133. Birtas-Atesoglu, E., et al., *Serum levels of free heat shock protein 70 and anti-HSP70 are elevated in Behcet's disease*. Clin Exp Rheumatol, 2008. **26**(4 Suppl 50): p. S96-8.
134. Koca, S.S., et al., *Anti-cyclic citrullinated peptide antibodies in rheumatoid arthritis and Behcet's disease*. Tohoku J Exp Med, 2007. **213**(4): p. 297-304.
135. Fresko, I., et al., *Anti-Saccharomyces cerevisiae antibodies (ASCA) in Behcet's syndrome*. Clin Exp Rheumatol, 2005. **23**(4 Suppl 38): p. S67-70.
136. Celet, B., et al., *Anti-alpha B-crystallin immunoreactivity in inflammatory nervous system diseases*. J Neurol, 2000. **247**(12): p. 935-9.
137. Tanaka, T., et al., *Behcet's disease and antibody titers to various heat-shock protein 60s*. Ocul Immunol Inflamm, 1999. **7**(2): p. 69-74.
138. Tasci, B., et al., *Humoral immune response to mycobacterial heat shock protein (hsp)65 in the cerebrospinal fluid of neuro-Behcet patients*. Clin Exp Immunol, 1998. **113**(1): p. 100-4.

139. Gheita, T.A., H. Samir, and H. Hussein, *Anti-annexin V antibodies in neuro-Behcet patients: clinical significance and relation to disease activity*. Int J Rheum Dis, 2012. **15**(5): p. e124-6.
140. Iaccarino, L., et al., *Anti-annexins autoantibodies: their role as biomarkers of autoimmune diseases*. Autoimmun Rev, 2011. **10**(9): p. 553-8.
141. Davatchi, F., et al., *Diagnostic value of pathergy test in Behcet's disease according to the change of incidence over the time*. Clin Rheumatol, 2011. **30**(9): p. 1151-5.
142. Ozdemir, M., et al., *Evaluation of application of multiple needle pricks on the pathergy reaction*. Int J Dermatol, 2008. **47**(4): p. 335-8.
143. Chang, H.K. and K.S. Cheon, *The clinical significance of a pathergy reaction in patients with Behcet's disease*. J Korean Med Sci, 2002. **17**(3): p. 371-4.
144. Dilsen, N., et al., *Comparative study of the skin pathergy test with blunt and sharp needles in Behcet's disease: confirmed specificity but decreased sensitivity with sharp needles*. Ann Rheum Dis, 1993. **52**(11): p. 823-5.
145. Akman-Demir, G., P. Serdaroglu, and B. Tasci, *Clinical patterns of neurological involvement in Behcet's disease: evaluation of 200 patients*. The Neuro-Behcet Study Group. Brain, 1999. **122** (Pt 11): p. 2171-82.
146. Ghasemi, Y., et al., *The possible role of HSPs on Behcet's disease: a bioinformatic approach*. Comput Biol Med, 2012. **42**(11): p. 1079-85.
147. Hirohata, S., *Histopathology of central nervous system lesions in Behcet's disease*. J Neurol Sci, 2008. **267**(1-2): p. 41-7.
148. Lamarca, V., et al., *Two isoforms of PSAP/MTCH1 share two proapoptotic domains and multiple internal signals for import into the mitochondrial outer membrane*. Am J Physiol Cell Physiol, 2007. **293**(4): p. C1347-61.
149. Xu, X., et al., *The novel presenilin-1-associated protein is a proapoptotic mitochondrial protein*. J Biol Chem, 2002. **277**(50): p. 48913-22.
150. Thurner, L., et al., *Progranulin antibodies in autoimmune diseases*. J Autoimmun, 2013. **42**: p. 29-38.