UNIVERSITY OF LATVIA

HUMORAL ANTIGENS IN CANCER: THE BASIS OF IMMUNOGENICITY AND RELEVANCE FOR CLINICAL APPLICATION

Doctoral thesis

Department of Molecular Biology and Biochemistry
FACULTY OF BIOLOGY

Author: Karīna Siliņa
Supervisor: Dr. biol. Aija Linē

Riga 2010

Great truth is a truth whose opposite
is also a great truth.

Niels Bohr

University of Latvia
Promotion Council
Biology
Riga, Latvia

Chairman: **Prof., Dr. biol. Viesturs Baumanis**
(University of Latvia, Riga, Latvia)

Opponents: **Prof., Dr. habil. biol. Pauls Pumpēns**
(University of Latvia, Riga, Latvia)

Prof., Dr. med. Robert C. Rees
(Nottingham Trent University, Nottingham, United Kingdom)

Dr. biol. Una Riekstiņa
(University of Latvia, Riga, Latvia)

Promotion paper is accepted for the commencement of the degree of Dr. biol. (in Molecular Biology).

Commencement date 21.01.2011.

Abstract

The exploration of humoral responses to cancer may reveal diagnostic and prognostic biomarkers and may help to discover potential immunotherapeutic targets. In a previous study aiming to identify a comprehensive set of antigens eliciting IgG antibody production in melanoma, breast, gastric and prostate cancer patients by using T7 phage display-based SEREX approach, we isolated over 1300 serum-reactive clones comprising (i) antigens translated in the natural open reading frames (ORF) including known cancer-testis (CT) antigens, antigens previously detected by conventional SEREX and uncharacterised antigens representing wild-type non-mutated proteins, (ii) several putatively novel genes and splice variants with an undefined ORF, as well as (iii) peptides translated in an unnatural ORF. In the current study we tried to determine whether any of the novel antigens are potentially applicable as targets for immunotherapy and to get an insight into the mechanisms of their immunogenicity by detecting the frequency of antigen-specific humoral responses in cancer patients and healthy controls, by analysing the expression pattern in various normal and cancerous tissues, as well as by characterising naturally occurring T cell responses for two immunotherapy candidate antigens.

Autoantibody profiling assay for melanoma, breast and gastric cancer patients and healthy donors were performed using our custom developed antigen microarrays. To select the most promising novel antigens recognised by cancer-related antibodies for further experimental analyses we established a prioritisation scheme where the antigens were evaluated and ranked by their frequency of cancer autoantibodies, EST profile, as well as functional and structural properties. The top 30 ranked natural and undefined ORF antigens were selected for further mRNA expression analysis, and revealed six annotated genes and four novel genes that were predominantly expressed in testis among various normal tissues. Two of these genes – LRRC50 and ESCO1, were significantly upregulated in cancers and represented novel CT antigens with potential use in immunotherapy. Two novel testis-associated splice variants were identified for CFL1, and COX6B2, and their upregulation was observed in some tumours suggesting that immunoprivileged tissue-specific splice sites can be recognised in cancer contributing to cancer antigenicity. 16 genes showed ubiquitous expression and four of them were significantly upregulated in tumours – ACTR2, LIG1, NOL8, and SPARC, however not exceeding the level in other normal tissues and the reasons of immunogenicity of these wild type proteins remains elusive.

Among the cancer-related natural ORF antigens was SPAG8 – a member of heterogeneous group of proteins called sperm associated antigens (SPAG) that are commonly expressed in male germ cells, are capable to elicit immune response underlying infertility, and several of them recently have been shown to be expressed in cancers and may have oncogenic properties. This prompted us to investigate further the expression pattern and humoral immune response against SPAG proteins. Several of SPAG genes previously proposed as targets for cancer immunotherapy were found to be ubiquitously expressed in normal human tissues, while SPAG1, SPAG6 and alternative splice variants of SPAG8, SPAG16 and SPAG17 were predominantly expressed in testis and elevated in various tumours with frequencies ranging from ~10% ~70%, further supporting the idea of testis-associated splice site recognition in cancer. The upregulation of SPAG6 in cancer was confirmed by immunohistochemical analysis of tissue microarrays. Cancer-associated spontaneous humoral immune response was detected against SPAG1, SPAG6, SPAG8 and a novel testis-restricted splice variant SPAG17-A1 and ascribe these genes as novel CT antigens that potentially are applicable as immunotherapeutic targets and serological biomarkers. In the case of SPAG17-A1 we demonstrated that the humoral response is

directed against the novel alternative testis-restricted splice variant and we propose to designate these antigens as CT-spliced antigens.

In order to establish a robust and reproducible assay for the analysis of spontaneous T cell responses, we participated in the international study aiming to standardize ELISPOT procedure what resulted in the generation of a reference protocol. Next, an assay for the isolation of CD4+ and CD8+ T cells from peripheral blood and *in vitro* pre-sensitisation of T cells was elaborated and tested with positive control viral peptides and well-known tumour antigen NY-ESO-1, and the results demonstrated that it allows the detection of tumour antigen specific CD8+ T cells in the peripheral blood mononuclear cells after *in vitro* pre-sensitisation. The initial CD8+ T cell activation analyses of two candidate antigens LRRC50 and SPAG8 showed a plausible SPAG8-specific naturally occurring CTL response in one gastric cancer patient suggesting that it could be possible to achieve also a therapy-induced response.

In conclusion we have characterised the mRNA expression and humoral immune response of novel cancer antigens revealing novel potential immunotherapy target genes, and antigens representing testis-associated alternative splice variants that could form a novel cancer antigen category called CT-spliced antigens.

Table of Contents

Introduction

Malignant disease has a severe impact on the overall public health substantially shortening the lifespan of the ill despite various medical advances and still is a heavy burden for sustainable healthcare development. Breast cancer is the most frequent tumour type in Latvia and its incidence has increased by around 10% in the last ten years with more than 1000 new cases per year (from Latvian Central Bureau of Statistics). Gastric cancer is the third most frequent type of malignancy in Latvia with around 700 new cases per year and the second most frequent reason of cancer mortality (from Latvian Cancer Register). Melanoma although with comparatively low incidence in Latvia among European countries (data of year 2006 from European Cancer Observatory) is one of the most aggressive tumours affecting all age groups of society and the incidence has also increased during the last ten years reaching over 180 cases in 2008 (from Latvian Cancer Register). The major treatments of these malignancies are surgical, hormone, radio or chemotherapies that significantly decrease patients' quality of life during treatment, and the curative outcome for certain tumour subtypes and late stages is rare.

One of the modalities of cancer treatment utilising the natural defence mechanisms of the human body called tumour immunotherapy holds great promise for the clearance of malignant cells due to highly specific targeting against tumour-specific and tumour-associated antigens, and for the possibility to substantially increase patients' quality of life due to low toxicity. The first line of clinical application of tumour immunotherapy has been the administration of stimulatory cytokines like interleukin 2 (IL-2) and interferon alfa (IFNα) as well as monoclonal antibodies against tumour antigens either conjugated to a therapeutic agent or with opsonising effect[1]. The application of the cellular compartment of the immune system in clinics, however, has not been so straightforward and has experienced a full spectrum of fluctuations from the disbelief in the capability of immune cells to control cancer to the first cellular therapy receiving FDA approval in spring 2010 marking a cornerstone in the development of adoptive immunotherapy. Hitherto a lot of effort has been put into various strategies of adoptive tumour immune response activation mainly divided into two different approaches depending on the used antigenic determinants. The activation of immune response can be achieved through either (i) tumour antigen-specific targeting by vaccination with antigen-encoding DNA, mRNA, peptide/protein, or by administration of *in vitro* propagated antigen-specific T cells, or (ii) against whole tumour, when the stimulant is tumour cells, dendritic cells pulsed with total tumour cell biomaterial, tumour cell HSP-peptide pool vaccines, or *in vitro* propagated tumour infiltrating lymphocytes (TILs)[1]. Either modality has struggled with (i) innate and/or adaptive suppression of the induced immune response by myeloid-derived suppressor cells, tolerogenic dendritic cells and inducible T regulatory cells, as well as suppressive cytokines produced by tumour or its stroma, and (ii) the selection of tumour escape variants by decreasing antigen presentation or losing antigen expression[2,3]. However, the combination of vaccines with adjuvants and blockers of tolerance, that skews the balance from tolerogenic to activating conditions, or with chemotherapy or radiotherapy that dampen the activity of suppressor cells and promote the immunogenicity of tumours by increasing antigen diversity and expression have shown promising trends[4,5].

Regardless of either antigen-specific or whole tumour targeting methods, the endpoint of a successful adaptive response is the expansion of cells specific to a cancer antigen capable of controlling and destroying the tumour with a higher curative probability if the response is polyclonal and against many antigens. A large number of tumour-specific and associated antigens has been identified by using B and T cell based approaches, nevertheless, the field still lacks sufficient knowledge of the true protective antigens as

well as tools for monitoring the therapy induced immune response, hence there is a continuous necessity for the identification and characterisation of cancer proteins that the immune system can recognise. A wide expert panel has recently elaborated guidelines of prioritising antigens for immunotherapy to discern the ones that could successfully be targeted in clinical trials and name the specificity and oncogenicity of a novel antigen as the dominant criteria for choosing it for further immunogenicity and therapeutic function evaluations[6]. Previously we applied a high-throughput serological antigen identification approach - the phage-display SEREX, to characterise humoral cancer immunome - a comprehensive set of proteins that are recognised by autoantibodies in sera of melanoma, breast, prostate and gastric cancer patients, which resulted in the collection of over 1300 antigenic clones. The aim of this doctoral study was to define whether any of the identified novel humoral cancer antigens may serve as immunotherapeutic targets by characterising their expression pattern in various normal tissues and looking for evidence of overexpression in tumours, as well as analysing the frequency of spontaneous humoral and cellular immune response in cancer patients.

1 Literature overview

1.1 The basis of adaptive immune response

The adaptive immune system has evolved as a means to improve the host's defences against all sorts of pathogens, and is based on the generation of cells with a plethora of receptors capable to bind different biological macromolecules called antigens that mediate the further protective effector functions[7]. The enormous receptor diversity is created through an intense genomic reorganisation of the corresponding genes, which is a "blind" process that inevitably creates not only receptors recognising molecules which are characteristic to the pathogen – the so-called "non-self" or foreign antigens, but also receptors recognising molecules of the host's body called "self" or autoantigens[8]. There are two major lineages of effector cells executing the adaptive protection namely B and T lymphocytes that underlie the humoral and cellular immunity possessing highly destructive potential capable of clearing large amounts of parasites, bacteria, virally infected cells or cancer cells. The activation of the adaptive immunity is essential for survival, while the tolerance mechanisms ensure that immune cells bearing self-recognising receptors won't be propagated and prevent severe damage to the host defined as autoimmune diseases[8]. The balance between activation and tolerance is kept by a milieu of regulator mechanisms that together underlie normal tissue homeostasis.

1.1.1 The ignorance of "self", establishment of immune tolerance

Both effector cell lineages derive from hematopoietic stem cells in the bone marrow but differ in the further maturation fates. The differentiation of B cells (here and further designating the B2 lineage lymphocytes) takes place in the bone marrow, whereas T cell precursors migrate and mature in the thymus, and the first and most stringent control of self-reacting receptors named the central tolerance takes place in these primary lymphoid organs[9]. The central tolerance is based on the selection of cells for either purging or further maturation by evaluating the binding affinity of both, developing T and B cells receptors named pre-BCR and pre-TCR to self-antigens presented in the primary lymphoid tissues. The mechanisms through which thymus presents peripheral tissue self-antigens during T-cell development include the expression of autoimmune regulator AIRE in thymic medullary epithelial cells that causes promiscuous expression of various genes or by the immigration of antigen-presenting immature dendritic cells[10]. Much less is known about the autoantigen presentation to developing B cells, but in both cases strong receptor binding leads to receptor editing, clonal deletion, or anergy[9,11] with receptor editing being the predominant mechanism for pre-B cells and the other two for pre-T cells[10]. The effects of intermediate avidity binding of the pre-receptors differ for pre-B and pre-T cells leading to receptor editing or positive selection, respectively, while no pre-receptor binding induces apoptosis due to the lack of survival signals for pre-T cells but ensures pass of pre-B cells[9,12]. Lymphocytes that have passed the central tolerance then continue their differentiation, and T cells depending on the TCR stimulation properties diverge into distinct subtypes like CD8+ cytotoxic T lymphocytes (CTL), CD4+ T helper (Th) cells, as well as so-called natural CD4+ T regulatory (nTreg) cells that evolve as a second scenario in case of a strong autoreactivity of the pre-TCR and possess strong suppressive activities against self-antigen specific lymphocytes in the periphery[10]. Cells that leave to periphery are from this point described as naïve or antigen-inexperienced[10].

Although very efficient, central tolerance cannot eliminate all self-reactive cells as not all self-proteins are displayed by the primary lymphoid organs as well as the positively selected T cells that bind self-antigens with intermediate affinity can possess a threat, hence further control mechanisms are applied after lymphocytes have exited the primary

lymphoid organs called the peripheral tolerance that further diminish the autoreactive repertoire, control the activation of naïve effectors as well as control the course of immune response[13].

The number of autoreactive B cells is further decreased through the negative selection of high avidity autoantigen-binding BCRs in the spleen where the transition from immature to mature B cells takes place[10]. This selection is additionally fine-tuned by integrating signals from the engaged BCR and antigen-independent receptor BAFFR that is crucial for B cell survival and differentiation[14], and introduces a certain degree of freedom in this censoring step resulting in about 20% of autoreactive B cells in the mature naïve repertoire in the periphery[12]. Also Th cells recognising the self-peptide-MHCII complexes presented on the stimulated B cells but not being activated themselves can induce apoptosis of B cells through Fas-FasL interactions[11], as well as by activated Treg cells[15]. During adoptive immune response B cells that receive proper activation from Th cells undergo somatic hypermutation or affinity maturation of their BCRs that can result not only in enhanced BCR affinity towards the foreign antigen eliciting the activation, but also in the generation of autoreactive receptors. At this point B cells that have acquired such a *de novo* self-reactivity are controlled by a distinct Th cell subset called the T follicular helper (T_{FH}) cells that are also crucial for mounting of proper B cell activation[16].

Peripheral T cell tolerance and the principles of self-non-self discrimination are described by the so-called "Danger model"[17]. In order for a T cell to bind its antigen, it must be firstly taken up, then intracellularly processed, and finally presented to the T cell on the major histocompatibility complex (MHC) II by antigen-presenting cells[18]. The most important cell with this capacity is the dendritic cell (DC), which collects the surrounding material while circulates in tissues and scans the environment for the so-called danger signals such as conserved pathogen molecules like lipopolysaccharides or tissue damage signals like HMGB1, DNA or purine metabolites (termed as pathogen-associated molecular patterns, PAMPs, and danger-associated molecular patterns, DAMPs, respectively), and if these signals are encountered by corresponding receptors (like Toll-like receptors), DCs differentiate from immature to mature state necessary for proper T cell activation[19]. DCs that present antigens collected from healthy tissue environment and haven't encountered any danger signals are in an immature or incomplete maturation state that is characterised by the lack of co-stimulatory molecules on their surface and no or poor stimulatory cytokine production. Immature DCs induce anergy or death in T cells recognising presented antigens[13] or induce their differentiation into regulatory subtypes like inducible T regulatory (iTreg) cells, which have the same censor functions as the nTreg cells[20].

Next line of control are the Treg cells that, if activated by antigen-presenting cells, inhibit both, effector CD4+ and CD8+ T cells in either contact-independent manner by secreting suppressive cytokines IL-10, TGFβ and IL-35, or in contact (or at least a very close proximity)-dependant manner by (i) modulation of antigen-presenting cells from activating to tolerogenic, (ii) induction of apoptosis of effector T cells through granzyme B, perforin, TRAIL or survival cytokine IL-2 deprivation, and by (iii) the metabolic disruption of effector T cells through the release of adenosine nucleosides or IDO by dendritic cells[21]. Lately the stromal cells of lymphoid organs like lymph node fibroblastic reticular cells[22] and lymph node-resident endothelial cells[23] have been shown to express peripheral tissue-restricted antigens in either AIRE-dependant or independent manner, respectively, and present them to CD8+ T cells inducing tolerance[24].

Besides Treg activities, several T cell control mechanisms take place in the periphery that are intrinsic to the effector cells: (i) repeated TCR activation causes activation-induced cell death (AICD) (this also contributes to T cell selection in thymus)[25], (ii) activating TCR

engagement for prolonged periods of time causes apoptosis or desensitisation of TCR signalling in CD4+ T cells termed adaptive tolerance[13], or functional exhaustion for CD8+ T cells[26], as well as (iii) TCR engagement abrogates initial proliferation of CD8+ T cells if no additional Th stimulation in terms of mitogenic cytokine IL-2 is received called antigen-induced non-responsiveness (AINR)[27].

Another "mechanical" way of peripheral tolerance is the mode of the migration of naïve lymphocytes that is restricted to the blood-lymph system, and the anatomical barriers such as blood-brain, eye, testis or placenta ensuring immunological ignorance by reduced level of adhesion molecules on draining blood vessels necessary for antigen-experienced lymphocyte exit to tissues, as well as in testis decreased MHCI expression and apoptotic ligand expression in case an attacking lymphocyte did arrive, and such tissues are termed as immunoprivileged[13,28].

Additional B and T lymphocyte subsets have been identified with inhibitory capacity like regulatory B cells secreting immunosuppressive cytokine IL-10[29] and regulatory CD8+ T cells (both, naturally occurring and inducible[30]) that kill or anergise activated T cells antigen-specifically or inhibit antigen presenting cells[31,32]. The latter has yielded a novel model describing the principles of peripheral self-tolerance named "Avidity Model of Self-Nonself Discrimination", which encompasses the idea that the immune system achieves self-non-self discrimination not by recognising the structural differences between self versus foreign antigens (as underlined by the "Danger model", where the maturation of dendritic cells is mounted as a result of PAMP or DAMP recognition), but rather by perceiving the avidity of TCR-antigen interaction[33]. Taking into account that there are no high avidity self-reacting T cells passing the central tolerance or Treg surveillance, the positively selected peripheral T cells may bear low or intermediate avidity TCRs to self antigens as well as low, intermediate or high avidity TCRs to foreign antigens. The model predicts that all low or intermediate avidity T cells are anergised or deleted in the periphery regardless to their self or foreign specificity, while all high avidity TCRs overcome the suppressive barrier, thus not only are the potentially autoimmune T cells downregulated, but also the strongest foreign antigen-specific T cells are selected enhancing the protective response[33]. This model is also supported by a recent review reporting that suboptimal engagement of the TCR by a variety of peptide-MHC ligands (such as antigens with intermediate avidity) triggers T-cell anergy[34].

1.1.2 The recognition of "non-self", mounting of protective immune response

According to the "Danger model" the initial key event of immune response activation is the induction of DC maturation by danger signals. Upon maturation DCs increase migration towards lymph nodes, upregulate co-stimulatory molecules and secrete a stimulating cytokine profile depending on the danger signal type[20]. The role of lymphoid organ stromal cells as antigen presenters and important regulators of immune cell trafficking and activation by chemokine and cytokine production have also been increasingly acknowledged[35] and together with mature DCs ensure all necessary signals for effector cell activation.

CD4+ or Th cells are the first lymphocytes that need to be activated to further properly mount both, the cellular cytotoxicity effector (CTL) and humoral effector (B cell) activation. To induce the proliferation and effector functions of a naïve Th cell three signalling checkpoints have to be fulfilled: it needs to bind to the complementary antigen in an MHCII context (compatible with the CD4 co-receptor) (signal 1) on a matured dendritic cell that expresses co-stimulatory molecules like B7-1/2 binding to stimulatory co-receptors on the Th cell like CD28 (signal 2) as well as produces stimulating cytokines (signal 3)[20]. After all three signals have been received, Th cell proliferates and

differentiates into various functionally different effector subtypes like Th1, Th2, Th17, follicular helper T (T_{FH}) cells etc. (Figure 1) depending on the received signal 3 cytokine profile as well as the strength and duration of TCR engagement[36]. These subtypes are epigenetically determined and can switch their phenotype to other subtypes in response to dynamically changing cytokine profile provided by antigen-presenting and regulatory cells to orchestrate the course of the immune response[37].

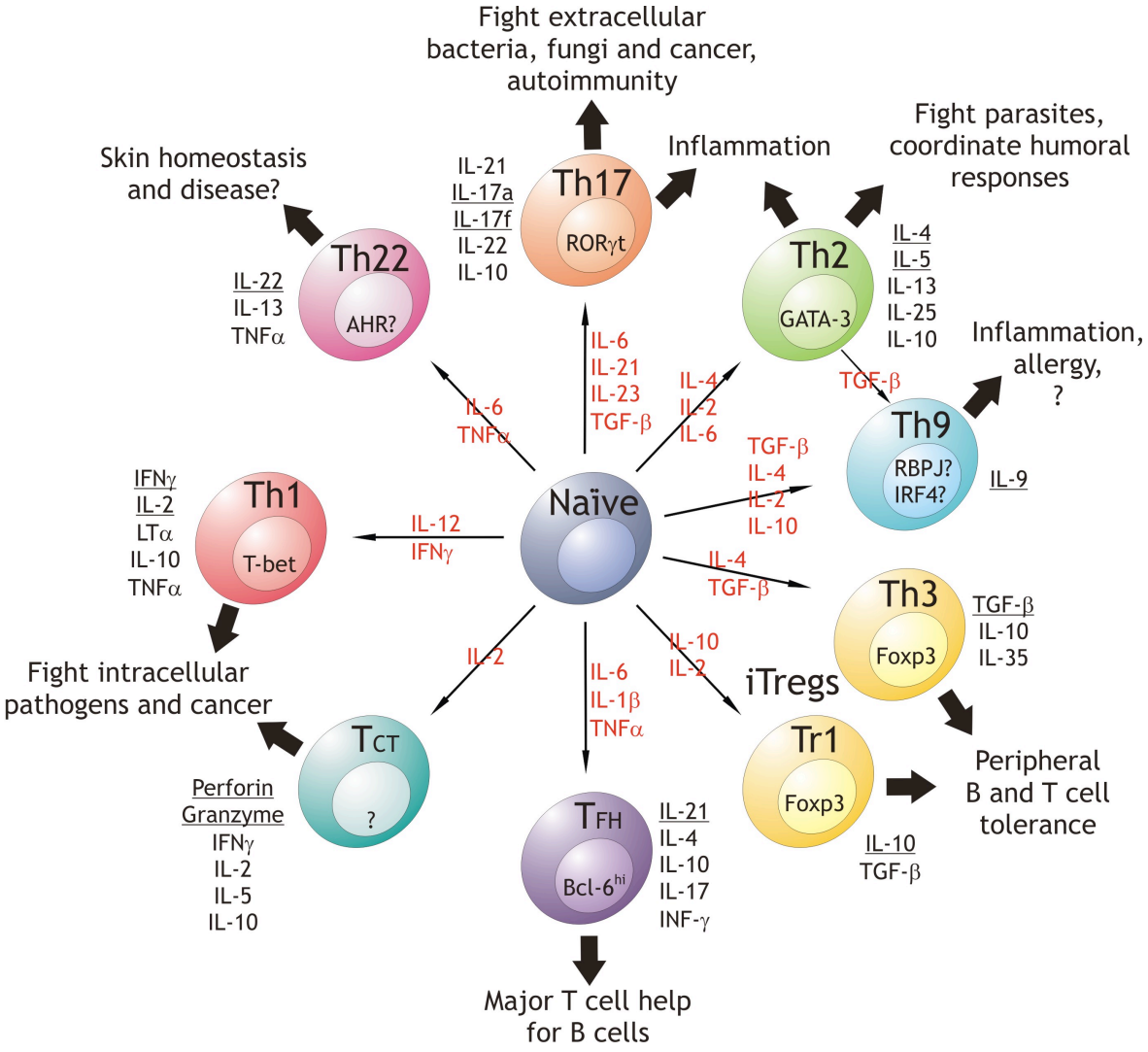


Figure 1. Summary of CD4+ T cell subset differentiation and functional implications.

The main currently known CD4+ T cell subtypes with lineage defining transcription factors and key functions are indicated. The major cytokines inducing certain differentiation paths are indicated in red, major secreted cytokines of effector cells are indicated in black and those that define functions characteristic to the particular T cell subset are underlined; adapted from Zhou et. al[36], Martorelli et. al[38], Zhu and Paul[39], and Annunziato and Romagnani[40].

Besides the well-established Th1, Th2 and iTregs, several other distinct Th cell subtypes have been identified. T follicular helper cells (T_{FH}) have surface CXCR5 chemokine receptors defining their migration to germinal centres, strong expression of Bcl-6 transcription factor leading to high levels of IL-21 production, and are the major subtype inducing B cell activation and differentiation into plasma cells during germinal centre reaction a role long ascribed to the Th2 cells[41-43]. A distinct lineage of T cells secreting high levels of IL-17 (Th17) was identified as important players in inflammation and autoimmunity[44], but lately also their importance in anti-cancer response has been described[45]. Recently effector T cells producing IL-9, but not IL-4 or IL-5, named Th9 cells have been described with implication in inflammatory responses, allergy and asthma[46], and the suggested but not yet approved characteristic transcription factors are RBPJ [47] and IRF4[48]. IL-22 producing cells that express skin-homing receptors and don't secrete IL-17 have been described as Th22 cells[49-51]. Currently they are suggested to play a role in early epidermal immunity and remodeling[52,53], and a potential transcription factor defining this subtype is aryl hydrocarbon receptor (AHR)[54]. Cytotoxic CD4+ T cells (T_{CT}) secreting cytolytic enzymes as perforin and granzyme have been described already more than two decades ago, but only in the last years has this capacity gained attention, and these cells have been shown to be important in fighting infections and cancer[38].

Antigen-inexperienced CD8+ T cells can be activated only if their TCR has encountered a compatible antigen presented in an MHCI (compatible with the CD8 co-receptor) context on matured DCs (called cross-presentation) that also provide analogous signals 2 (co-stimulatory surface molecules) and 3 (cytokines). For a dendritic cell to provide the necessary activating signals to the CTL, it needs to be conditioned by an activated Th cell to produce the proper cytokines like IL-12 and IFNα[27]. Th cells can also produce cytokines directly providing signal 3 to naïve CTLs, for example IL-21 that is mostly produced by activated Th17[39] or T_{FH} cells[55] is found to be a potential alternative signal 3 cytokine capable of inducing naïve CD8+ T cell activation and differentiation into highly cytolytic effectors[56]. Other cytokines produced by Th cells can enhance or modify naïve CTL activation and induce differentiation into several effector subtypes (Figure 2), for example Th2 cytokine IL-4 skews the differentiation to a Tc2 effector subtype that produces IL-4 (so-called type 2 cytokines and type 2 response), while Th1 cytokines IL-2 and IFNγ promote conventional CTL or Tc1 subtype differentiation and clonal expansion (so-called type 1 cytokines and type 1 response)[27]. After activation effector CTLs migrate out of the lymphatics following the signals of tissue damage and inflammation and search for cells presenting the specific antigen on their MHCI complexes and once encountered induce their apoptosis by direct (Fas-FasL or granzyme B and perforin) or indirect (TNFα, IFNγ) mechanisms[57].

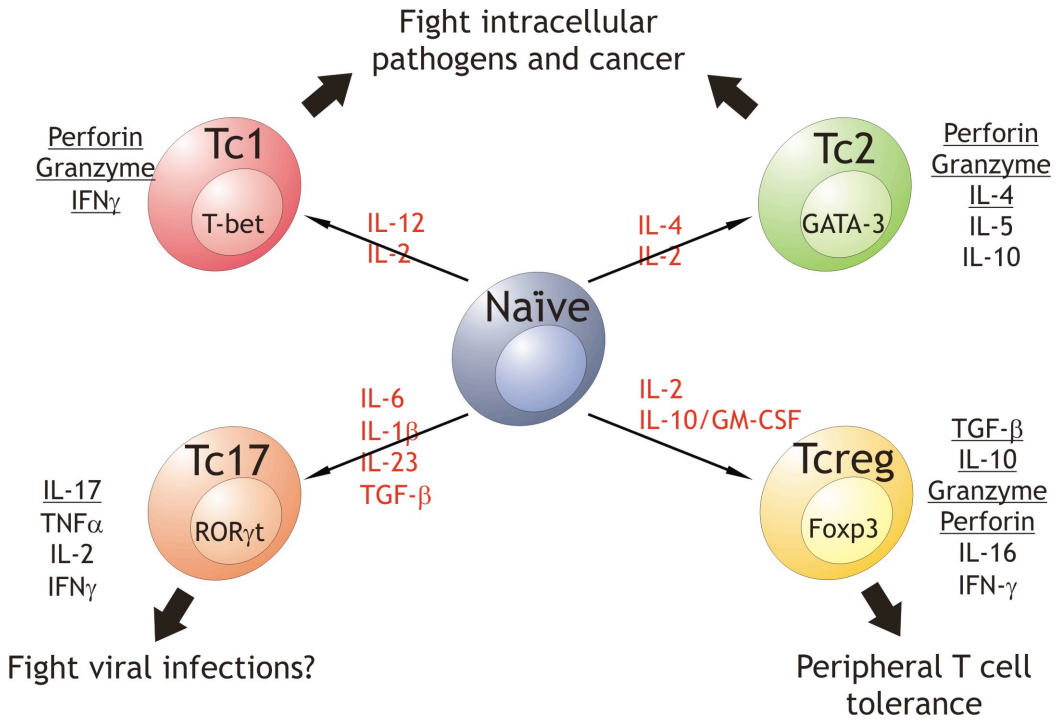


Figure 2. Summary of CD8+ T cell subset differentiation and functional implications.
The main currently known CD8+ T cell subtypes with lineage defining transcription factors and key functions are indicated. The major cytokines inducing certain differentiation paths are indicated in red, major secreted cytokines of effector cells are indicated in black and those that define functions characteristic to the particular T cell subset are underlined.
The existence of similar polarized subsets of CD8+ T cells as for Th cells is known already for over two decades[58]. These are equally cytolytic, but Tc1 cells show more efficient cytotoxicity both in autoimmunity[59] as well as cancer[60]. Although first described nearly 40 ears ago[61], only recently a regulatory subtype of CD8+ T cells (Tcregs) has increasingly been studied and revealed several subtypes of Tcregs with a variety of proposed regulatory actions reviewed by Suzuki et. al[30]. Last couple of years have yielded another distinct CD8+ T cell subtype analogous to the Th17 cells named Tc17 that have poor or lack cytolytic activity. Although their functions have not been clearly established, they have been shown to be important in the protection against lethal infection of influenza virus in mouse models[62,63].

The major activation checkpoints necessary for high affinity antibody-producing B cell generation also follow the sequential signal paradigm[64]. The presence of (1) corresponding BCR ligand (the antigen) that is internalized after binding to the BCR and presented in the context of MHCII to Th cells, and (2) activated T helper cells bearing TCRs specific for the presented internalized antigen and producing stimulatory cytokines ensures cognate activation of effector B cells, and their further differentiation fate is determined by the kind and dose of the antigen, the location of the encounter and by the combination of available cytokines yielding either cells with or without somatically hypermutated BCRs (following follicular or extrafollicular pathways, respectively), short or long-lived antibody secreting plasma cells or memory B cells[64-66]. Other subtypes of B cells like marginal zone B cells of the spleen can produce low affinity antibodies without T cell help and ensure a first line of defence and rapid response to blood born antigens[65]. The antigens are brought to naïve T cells by the DCs, while B cells can encounter antigens diffused in the blood or lymphatics. Lately also lymphoid organ cells as subcapsular sinus macrophages, follicular dendritic cells, as well as marginal zone B cells and B cells them selves have been shown as capturers and concentrators of immune complexes containing antigen-bound natural or early response low affinity antibodies together with complement components in the lymph nodes and spleen to facilitate their recognition by antigen-specific B cells[67]. The major Th cell subset to provide the second signal is the T_{FH} cells that provide essential stimulation for the induction of receptor affinity maturation and antibody class switching through cell-cell contacts as well as cytokines like IL-21, and ensure the selection of the fittest BCRs after somatic hypermutation[55]. Cytokines secreted by other Th cell subtypes as Th1 and Th2 also influence the functional properties of B cells like the choice of produced class switched antibodies as well as B cell secreted cytokine profile that corresponds to the type 1 and 2 response, respectively[68]. Together with cytokine secretion, B cells can reciprocally influence the choice of Th subtype differentiation due to their antigen-presenting functions providing additional stimulation to T cells[69]. The most important effector function of B cells, however, is the production of antigen-specific antibodies, which exert various protective effects like inactivation of soluble toxic products of pathogens or blocking of pathogen surface receptors necessary for infection, induction of antibody dependant cellular cytotoxicity (ADCC) or complement dependant cytotoxicity (CDC), as well as opsonisation (marking the bound cells for phagocytotic destruction by the innate immunity cells like macrophages, neutrophils, etc.), as well as promote the uptake of the immune complexes by antigen presenting cells facilitating further B and T cell activation[70]. The cytotoxic outcome of antibody binding through ADCC is ensured by cytotoxic immune cells (such as CTL as well as cells of the innate immunity like natural killers (NK), natural killer T cells (NKT), etc.) that carry receptors for antibody heavy chain constant regions like Fcγ for IgG class antibodies and induce their lytic activities upon receptor stimulation in contact with the antibody-covered cells[71]. CDC is underlined by the capability of the complement component C1q to bind to the constant regions of antibodies covering a target cell and induce further activation of the complement system resulting in membrane attack complex formation and target cell lysis; alternatively, complement-bound antibodies can be recognised by immune cells possessing complement receptors that can lyse or phagocytose the covered cells[71].

1.2 Adaptive anticancer immune response

1.2.1 Cancer immune recognition, the immunosurveillance

Cancer cells are characterised by considerable genetic instability that impacts the produced proteins both in a qualitative and quantitative way. Mutations, chromosomal

translocations and deregulation of alternative splicing, translation, and post-translational modifications (PTMs) can all create structurally altered proteins that have not been centrally tolerated. As well changes in gene expression regulation by epigenetic aberrations, microRNA deregulation, and decreased protein degradation cause elevated levels of proteins otherwise present in minor amounts in normal cells leading to increased presentation on MHCI of tumour cells as well as increased release in tissue fluids due to tumour cell death leading to increased uptake and presentation by DCs what can surpass the peripheral tolerance barriers. The immune system can recognise such "altered-self" proteins or cancer antigens and mount adaptive response targeted against the cells containing them[2].

According to the Danger model, the dendritic cell has to receive signals reporting about stress in the environment in order to be able to activate protective adaptive humoral or cellular immune response. Such danger signals in the evolving tumour microenvironment are various molecules released from dying tumour cells due to hypoxia and unbearable genetic instability causing death by mitotic catastrophe, apoptosis or necrosis and thus delivering a message of tissue damage through such released molecules as calreticulin, uric acid, high-mobility group box 1 (HMGB1), heat shock proteins, etc.[72]. Stress signals are also sensed by the innate immunity cells like tissue residing macrophages, mast cells and other that release pro-inflammatory cytokines, tissue remodelling enzymes, reactive oxygen species, histamine etc. additionally harming tumour cells and stimulating the maturation of dendritic cells, tumour antigen presentation and immigration of lymphocytes to the tumour site together inducing the first protective response, the acute inflammation[73]. Many studies have underlined the importance of both, innate and adaptive immunity in recognising and eliminating cancer cells every day[74], and the presence of activated adaptive immunity cells in the tumour microenvironment have been shown to be a strong positive prognostic factor with a higher predictive capacity than any current markers routinely used in clinics[75].

The effector mechanisms dealing with tumour cell eradication are the direct cell contact-dependant or indirect soluble factor-dependant tumour cell killing by (i) cellular immunity effectors like CTLs of both Tc1 and Tc2 subtypes[76] as well as NKs and NKTs among the innate effectors[77], and by (ii) humoral immunity through antibody-mediated blocking of survival receptors on tumour cells, opsonisation, ADCC and CDC mechanisms[71]. Additionally cytokines produced by activated immune cells like INFγ induce increased MHCI expression and antigen presentation by tumour cells facilitating their further recognition[73].

1.2.2 The "altered-self", cancer antigens and reasons of immunogenicity

The adaptive immunity effectors need to meet the corresponding antigens that activated them in the lymphoid organs on tumour cells to be able to exert their protective activity. The realisation of the concept of tumour antigens that are specifically recognised by the immune system causing tumour regression led to idea to stimulate the immune response in a target dependant manner[78], however the field lacked substantial progress until Thierry Boon's group developed an approach to identify cancer antigens recognised by T cells, the pioneer being melanoma-associated antigen-1 MAGEA1[79]. Many studies since then have made a great effort in identification of such tumour-related targets utilising both, B cell and T cell-dependant strategies and yielded a large number of tumour antigens recognised by the adaptive immunity[80].

There are several classification approaches of cancer antigens. One way is by their molecular type, comprising protein, carbohydrate, lipid, glycolipid or glycoprotein antigens[81]. The antigens of non-protein nature are less immunogenic than their protein

counterparts as they are merely B cell or innate immunity antigens due to the capacity of MHCI and MHCII to display mainly peptide molecules[82]. However they represent a substantial part of altered phenotype of tumours and are important in cancer immunosurveillance[83]. The shift from the previous dogma of one-antigen-one-antibody to the realisation of the polyreactivity of antibodies with various epitopes as long as they fit into the pockets or corresponding paratope also called cross-reactivity or mimotope recognition, meaning that a peptide is able to mimic an epitope of non-protein or protein nature, has provided a novel approach to enhance the immunogenicity of such antigens by using synthetic peptide mimics conjugated to immunogenic carriers, and make the basis of the so-called mimotope therapy[84].

Antigens are most often classified by their specificity to cancer cells. The group of tumour-specific antigens can be further divided into (i) foreign antigens encompassing viral proteins in cases of virus-induced tumours or (ii) structurally altered self proteins (also called neo-antigens) that are formed due to the abovementioned mechanisms of genetic instability and deregulation of post-transcriptional events that create unique epitopes not met in normal cells. The second group called tumour-associated antigens are structurally unaltered self proteins with subcategories of (i) overexpressed antigens encompassing proteins that are ubiquitous but profoundly upregulated in tumours, (ii) cancer-immunoprivileged tissue antigens representing proteins present only in immunologically ignored tissues like germ cell development organs (commonly referred to as cancer-testis, CT antigens), retina, brain, placenta etc., (iii) differentiation antigens comprising terminally differentiated peripheral tissue proteins, and (iv) oncofoetal antigens[85,86].

Another way of antigen division is into conventional antigens translated from primary or natural open reading frames (ORF) of the corresponding genes and unconventional or cryptic peptide antigens encoded by introns of genes, exon-intron junction regions, regions created by chromosomal translocations or splicing defects, peptides translated from alternative ORFs (frame 2 or 3 of the coding region of mRNA) or subdominant ORFs (located in the untranslated regions of mRNAs)[80] or protein splicing[87]. The conventional antigens encompass the abovementioned tumour-associated antigens as well as part of the structurally altered tumour-specific antigens like the ones with missense mutations and altered post translational modifications, while the unconventional antigens are most often tumour-specific with an exception of deregulated alternative splicing if the created splice variant is characteristic to specific tissues and not a product of cryptic and/or non-canonical splice site recognition[88]. The immunogenicity of tumour-specific antigens is considered to be potentially higher than for the tumour-associated antigens due to the lack of any TCR-based tolerance, and protective antigens have been identified among them[1], however, one of the most immunogenic proteins to date is the CT antigen NY-ESO-1 (CTAG1B)[89] showing that tumour-associated antigens can also be potent activators of anti-cancer immune response[90].

1.2.3 Tumour evasion, the immunoescape

The current paradigm of tumour-immune system mutual interaction predicts the existence of three sequential phases: protection or elimination of tumour, an equilibrium state, and tumour escape together termed as cancer immunoediting[74]. The onset of clinically detectable cancer means that cells have acquired all six physiological properties or six hallmarks of cancer necessary for uncontrollable spread of tumour cells including self-sufficiency of growth signals, insensitivity to growth inhibitory signals, evasion from apoptosis, limitless replicative potential, induction of angiogenesis and capability to invade surrounding tissues and metastasize[91]. The appreciation of the role of immune system in

the control of cancer has introduced the capability of tumours to escape from immune protection as the 7th hallmark of cancer[92]. There are many strategies utilized by tumours to evade immune response affecting again both, innate and adaptive mechanisms.

The equilibrium phase is described as a latent period during which the tumour is not eradicated but is kept under control[74]. This phase is thought to be the time when tumours can develop escape phenotypes in response to the immunological pressure and implies the immune system as a tumour-promoting factor[77]. The selective pressure can result in clonal expansion of less immunogenic cells through two processes differing conceptually in the mechanisms ensuring the eventual escape called immunoselection that implies avoidance or hiding from immune response, and immunosubversion implying active suppression of immune response[74]. Immunoselection strategies used by tumour to become invisible or insensitive to immune recognition include (i) alterations of recognised epitopes (antigenic drift)[93], (ii) decreased expression of recognised antigens, (iii) diminishing antigen presentation, and (iv) downregulating death receptor signaling[94]. Immunosubversion is defined by the capability of tumour to suppress immune response either systemically (i) by facilitating the development of tumour antigen-specific Treg cells through secretion of the same suppressive cytokines[95], (ii) by promoting the development of myeloid derived suppressor cells[96] through tumour exosome secretion[97] that further orchestrate the downmodulation of DCs and adoptive effectors[98], or locally (iii) by expressing apoptotic ligands directly inhibiting immune cells, (iv) through production of suppressive cytokines as TGFβ and inhibitors as IDO, nitric oxide and arginase-1 that impair the maturation of dendritic cells[2], and (iii) by promoting innate inflammatory responses[75,99-101]. Indeed, while the acute inflammation is crucial for the induction of protective immune response, a large body of evidence shows that, if uncontrolled and turning into chronic states, inflammation has profound tumour promoting abilities participating in every aspect of tumour development and that elevated risk of tumour incidence is observed in patients with various chronic inflammatory conditions of various aetiologies including persistent infection, environmental irritants, or genetic alterations in genes orchestrating the resolution of immune response[99]. The low oxygen tension response is one of important oncogenic factors in chronic inflammation when macrophages upregulate hypoxia-inducible factors (HIFs) 1 and 2 that promote neovascularisation at the oxygen deprived sites inducing angiogenesis[102]. The released free radicals promote DNA damage and exacerbate the genetic instability facilitating the selection of escaping tumour cells as well as suppress the adaptive effector cell activity, while secreted tissue remodelling factors as a response to tissue damage promote the metastatic potential of tumour cells[73].

Besides the chronic inflammatory responses, also the local dominant subtype of activated adoptive cellular and humoral effectors has been shown to be either protecting or tumour-promoting with the polarisation between type 1 and type 2 responses contributing, respectively[103,104]. Recently also the role of tumour stroma as a promoter of inflammation and tumourigenic environment has been acknowledged[105]. Another viewpoint of the reasons of tumour escape is the poor immunogenicity of cancer stem cells that are believed to derive from normal stem cells, which are known to be immunosuppressive and poor self-antigen presenters also termed as immunoprivileged[28]. These tumour-intrinsic together with stroma and immune system-intrinsic properties peak at the onset of the escape phase when clinically detectable tumours are formed, which often are refractory to therapy due to the stringent immune-mediated selection process, and at this point it is very rare that a "spontaneous" tumour regression is observed[74].

1.3 Current adaptive tumour immunotherapy

1.3.1 Modes of adaptive immunotherapy

The induced activation of protective cancer-specific adaptive immunity has been the hope of oncologists and immunologist for over two decades, and recently the first milestone success has been reached with the first FDA approved adoptive cell transfer regimen called "Provenge" for prostate cancer treatment. The road has been long and winding, and there is still a lot of unresolved questions regarding the induction and sustaining of protective immune response by therapy.

The adoptive immunotherapy approaches can be classified into two major groups depending on the used target stimulant comprising antigen-specific and undefined antigen-based therapies[106]. The concept of undefined antigen therapy is based on the application of the whole tumour immunome (all antigens in a tumour that can be recognised) for the induction of immune response without the knowledge of the actual protecting antigens, hence letting the immune system to choose the most appropriate targets, and the only determinant of therapy success is the evaluation of clinical development[106]. Such therapies include (i) whole tumour cell vaccines based on the administration of tumour cells that are modified to be immunogenic rather than tolerogenic[107], (ii) heat-shock protein (HSP)-based vaccines that utilize the capability of HSP to serve as the "danger signal" for DC activation and form complexes with tumour peptides hence delivering cancer antigens in an immunostimulating fashion[108], (iii) DC-based vaccines encompassing various whole tumour antigen delivery ways like loading of DCs with whole tumour peptides, proteins, mRNA, DNA, whole tumour cell lysates, and necrotic or apoptotic tumour cells[109], as well as (iv) administration of *in vitro* propagated autologous tumour infiltrating lymphocytes (TILs)[1]. While such approaches can readily induce a potent adaptive immune response with high specificity against tumour, it is impossible to assess the immunological events that underlie either protection or failure to benefit without the information about the recognised antigen[110]. A second modality of adoptive immunotherapy that overcomes this limit utilizes a specific tumour antigen for the activation of anticancer response and allows to monitor the therapy induced immunological effects, which is crucial for the understanding of protective mechanisms *in vivo*[110]. Again various approaches have been developed and comprise vaccination with (i) antigen peptides, proteins, naked or incorporated in viral vectors antigen encoding RNA or DNA, (ii) DCs transfected with antigen encoding viral vectors, (iii) microparticle carriers conjugated with the antigen[106], as well as administration of (iv) *in vitro* expanded antigen-specific T cells[111]. Another antigen-specific targeting approach is by administration of monoclonal antibodies against the antigen, which either have a direct therapeutic effect (and such antibodies have successfully been used in clinics for already a couple of decades) or as anti-idiotype antibody-based vaccines. The latter also uses monoclonal antibodies specific to an antigen, but with a purpose to induce autologous antibody formation called anti-idiotype antibodies against the antigen-binding variable region of the administered monoclonal antibody, and if the paratope of the formed anti-idiotype antibody falls within the paratope of the original monoclonal antibody, it can mimic the original antigen and serve as a surrogate antigen to stimulate immune response[106].

The necessity to induce a specifically cancer-targeted immune response as well as the attractive opportunity to evaluate the course of immune response by assessing the dynamics of antigen-specific T cells has prompted the identification of a large number of tumour antigens, and guidelines for prioritisation of candidate antigens have been elaborated by introducing various evaluation criteria that each have a certain power or importance in the established rating system in order to discern those with the highest

potential to be successfully used in cancer immunotherapy[6]. The requirements named in the order of importance for the "ideal" cancer antigen as a therapy target are: efficacy in immunotherapy, immunogenicity, specificity to cancer cells, oncogenicity, high expression in all cancer cells including putative cancer stem cells, sharing among tumours of many patients, possession of multiple epitopes, location on cell surface or efficient presentation on MHC[6], and the search for this "ideal" antigen is not yet over if it ever will.

1.3.2 Major problems and possible solutions

The obstacles met by any type of initially developed immunotherapy approaches are the above described tumour-intrinsic and extrinsic (stromal and immune system-dependant) factors that initially caused fundamental doubt whether anticancer adaptive immunotherapy can be developed at all. For example, several studies have shown that, despite the activation and expansion of antigen specific CD8+ effector T cells, no clinical benefit was observed[112]. However, the acquired knowledge of immune escape mechanisms has now yielded additional weapons for immunologists in this battle. Both, the antigen-specific and unknown antigen-targeting approaches can be improved by combinations with (i) attenuators of the activity of regulatory cells[113,114] and the inhibitors of their chemotaxis and secreted suppressive molecules[115], (ii) adjuvants enhancing DC maturation and antigen cross-presentation[116,117], (iii) chemo and radio therapy that increases the death, antigen diversity and presentation of tumour cells and ablates myeloid suppressor cells[5,118], (iv) chemokines promoting effector cell trafficking to tumour sites[119], (v) co-stimulatory molecules to skew the activation of effector subtypes to Th1 type[120], as well as (vi) utilising different modes of antigen delivery[121], to name the most obvious.

One of the most successful examples is the addition of myeloablating total-body irradiation to lymphodepleting regimen prior to *in vitro* expanded adoptive tumour infiltrating T lymphocyte transfer of metastatic melanoma patients, which increased the response rates from 49% to 72% putting this combination as the most effective therapy of these patients refractory to other treatments[122]. And the currently first FDA approved adoptive cell transfer regimen "Provenge" is based on *ex vivo* processed dendritic cells expressing a key tumour antigen, prostate acid phosphatase, showing a statistically significant survival extension of at least 4 months in late-stage randomized trials increasing the life span for about 20% of these patients refractory to any other treatment[123].

Additionally, recent antigen-specific therapy advancements have been the generation of T cells with genetically engineered TCRs specific to any antigen of choice as a means to overcome difficulties related to the *in vitro* propagation of a few antigen-specific T cells till therapeutic numbers as well as increase the specificity and efficacy of TCR signalling, and have shown promising results[124,125].

2 Materials and Methods

2.1 Biological samples

2.1.1 Tissue samples

Tissues of melanoma, gastric, colon, and breast cancer and adjacent normal samples were obtained from patients undergoing surgery in Latvian Oncology Centre and stored in RNALater® (Applied Biosystems, USA). Samples of lung tumour tissues were obtained during standard diagnostic bronchoscopy procedure at Pauls Stradins Clinical University Hospital and stored in RNALater®. All collected samples were stored at -20°C before further processing. All patients have signed an informed consent and the study has been approved by the Central Commission of Medical Ethics of Latvia. The information about the collected biological samples for each patient together with clinical data is shown in Table 1.

Paraffin embedded tissue microarrays (TMAs) comprising duplicates of 45 various normal tissues or triplicates of 33 various normal tissues were purchased from Accu Max Array, USA (A700 (III)) and US Biomax, Inc., USA (FDA994), respectively. TMAs comprising single 40 various melanoma and 8 normal skin samples, and single 51 gastric cancer, 17 metaplastic, dysplastic and gastric inflammatory samples and 12 normal stomach samples were purchased from US Biomax, Inc., USA (ME481t and ST805t, respectively). TMAs comprising paired tumour and normal samples of 12 lung and 12 breast cancer patients were purchased from Accu Max Array, USA (A716(III) and A712(13), respectively).

2.1.2 Serum samples

Serum samples of melanoma, lymphocytic leukaemia, gastric, lung, colon, breast and thyroid cancer patients as well as healthy donors were provided by the Genome Database of the Latvian Population. Additional melanoma, gastric and prostate cancer patients' sera were kindly provided by Skin Cancer Unit in German Cancer Research Centre, Norwegian Radium Hospital and Onyvax Vaccine Therapies Ltd, UK, and Clinic of Gastroenterology, Hepatology and Infectious Diseases, Otto-von-Guericke University Magdeburg, Germany. The serum collection comprises 39 breast, 24 lung, 33 colon, 28 thyroid, 173 gastric (stage I – 32, II – 20, III – 31, IV – 37, not determined – 53), 52 prostate cancer, 190 melanoma (stage I – 22, II – 23, III – 21, IV – 22, not determined – 102), and 28 lymphocytic leukaemia patients as well as 126 patients of gastrointestinal inflammatory diseases (gastric ulcer – 17, duodenum ulcer - 20, gastritis and duodenitis – 10, acute hemorrhagic gastritis – 13, chronic atrophic gastritis – 48, dyspepsia – 11, Crohn's Disease – 8) and 153 healthy donors with no history of cancer or autoimmune disorders, infertility and chronic inflammatory diseases during last two years. All serum samples were stored at -80°C.

2.1.3 PBMC samples

Peripheral blood mononuclear cells (PBMCs) were obtained from 10 gastric cancer, 9 breast cancer patients and 5 healthy donors by Ficoll density gradient centrifugation using BD Vacutainer® CPT™ Cell Preparation Tubes (BD Biosciences, USA) containing sodium citrate as anticoagulant, Ficoll Hypaque solution and a gel layer separating Ficoll from the blood sample (Figure 3) according to manufacturers protocol. Briefly, approximately 8ml venous blood was drawn in each CPT™ vacutainer, using two vacutainers per individual. Blood samples were stored in room temperature in dark with gentle agitation until further preparation. No more than 5h after blood collection the tubes were centrifuged at 1700g for 20 minutes to separate peripheral mononuclear cells from all

other blood cellular constituents ensured by the Ficoll density gradient solution. Immediately after centrifugation the vacutainers were sterilized with ethanol and further processed in a sterile environment using sterile materials. The density gradient centrifugation causes the separation of blood in three layers – the top plasma layer, the middle PBMC and platelets layers and the bottom erythrocyte and granulocyte layer (Figure 1). Using a 3.5ml transfer pipette the middle layer is carefully removed and placed into a separate 50ml tube for each vacutainer. Collected PBMCs were washed as follows: warm (37°C) PBS (SIGMA-Aldrich, Germany) solution was added to the cells up to the 30ml mark, and tubes centrifuged 10 minutes 330g. The PBS was then removed not disturbing the cell pellet, and cells were immediately resuspended in the remaining drops by gentle taping not to let them sit in the pellet, and this was done always following centrifugation. The second washing step was done using warm 5ml RPMI1640 medium (Invitrogen, Norway) with 1x CTL Wash Supplement (CTL Ltd., USA) for each washing tube. At this point all tubes of the same donor were pooled together and number of cells counted under UV microscope using ethidium bromide/acridine orange (EB/AO) (SIGMA-Aldrich, Germany) staining to evaluate the proportion of live and dead cells. Cells were centrifuged as in the first wash and followed by slow resuspension in warm CryoABC media (CTL Ltd., USA) containing DMSO as the antifreeze according to the manufacturers protocol with adjusted cell concentration to 5-10x10^6cells/ml (usually around 2ml of final cell suspension per donor). Cells were then sampled in 1ml per cryovial tube and placed into Naglene® cooling tray that ensures slow temperature dropdown during the next 24h in -80°C. After 24h or not later than after 72h the cryovials were transferred to liquid nitrogen (liquid phase) for indefinite storage.

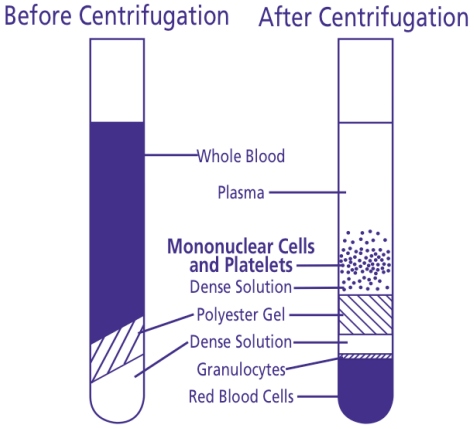


Figure 3. Separation of PBMCs from total blood by Ficoll density gradient centrifugation using BD Vacutainer® CPT™.

2.2 *In silico* analyses of humoral cancer antigens

Previously we applied phage-display SEREX to define a repertoire of antigens eliciting humoral responses in cancer patients that resulted in the identification of 1328 different phage clones representing 1158 non-redundant antigens as described in original paper I.

All of the identified antigens were entered in our custom antigen database that has built-in direct links to (i) NCBI nucleotide BLAST, to determine the identity of the clone, (ii) ExPASy Proteomics server (http://www.expasy.org/tools/dna.html) to obtain the amino acid sequence of the phage-displayed peptide, and (iii) NCBI protein BLAST, to determine the identity of the peptide. The antigens were then classified according to the translated peptide into three main categories including natural ORF antigens, unnatural ORF antigens comprising clones translated out-of-frame, and undefined ORF antigens comprising clones

that represent novel genes for which the ORF has not been characterised or novel splice variants representing alternative ORFs if there was any evidence that such ORFs can be generated as determined by (i) translation of predicted alternative exon combinations using Spidey tool (http://www.ncbi.nlm.nih.gov/spidey/), or by (ii) predictions of putative internal ribosome entry sites (IRES) using online tool IRSS (http://140.135.61.9/ires/) developed by Wu et. al[126] upstream of the possible translation initiation codon.

A built-in application was introduced to analyse EST tissue distribution. The EST sequences with E-value <1e-30, >70% coverage and >97% identities with an antigen clone were downloaded from ftp://ftp.ncbi.nih.gov/genomes/H_sapiens/; repeat sequences were masked by applying filters for low complexity regions and human specific repeats; ESTs were classified as "cancer" (Ca), "cell line" (CL), "normal" (N), "embryonic" (E), "germ cell" (GC) and "brain" (B) by extracting tissue source information from their annotations. The ESTs frequency of each antigen clone per category was calculated relatively to the total number of ESTs for that antigen normalized against the number of EST per each category and a graphical output of relative EST distribution among the categories is generated indicating the total number of EST in each category.

Also information about the predicted protein structure motifs and coiled-coil domains was entered for natural ORF antigens, which was obtained by using MOTIF Search (http://motif.genome.jp/) and http://www.russell.embl-heidelberg.de online software tools.

2.3 Analysis of humoral immune response

2.3.1 Screening of phage-displayed antigen microarrays

The autoantibody responses were analysed using our custom developed phage-displayed antigen microarray. The microarray production and screening were done as described in original papers I and III. Briefly, all identified recombinant antigen phage clones were amplified, purified and printed in duplicates on nitrocellulose-coated slides (Whatman, USA) to create the microarray that was screened with 1:200 diluted sera from 190 melanoma, 173 gastric, and 13 breast cancer patients, as well as 153 healthy individuals.

Besides the identified antigen collection, T7Select Phage Display vector (Novagen, USA) was used to clone full or partial reading frames of sperm associated antigens SPAG1, SPAG6, SPAG8, SPAG9, SPAG16 and SPAG17 including alternative splice variants. StrepII tag (Trp-Ser-His-Pro-Gln-Phe-Glu-Lys) was inserted downstream of the cDNA cloning site in the vector DNA to monitor the copy number of the recombinant proteins on each phage. The obtained recombinant phages were used for the production of SPAG antigen microarray that was screened with sera from 167 melanoma, 172 gastric, 52 prostate, 29 breast, 24 lung, 33 colon, 28 thyroid cancer, and 28 lymphocytic leukaemia patients, 126 patients with gastrointestinal inflammatory diseases as well as 147 healthy individuals.

To monitor the amount of phage per printed spot mouse anti-T7 phage tail antibody (Novagen, USA) was used. The serum reactivity was detected by Cy5 conjugated goat anti-human IgG antibody (Jackson ImmunoResearch, USA) and the ratio against the total amount of printed phage, which was detected by Cy3 conjugated goat anti-mouse IgG antibody (Jackson ImmunoResearch, USA) was calculated. Microarray data processing was performed elsewhere (Zayakin P, et. al, 2010, article in preparation). Briefly, a two-step normalisation strategy was used for the fluorescent signal ratios in order to eliminate variations introduced by custom production of microarrays and variable background intensities of different sera. At first, the values in each slide (each serum) were normalized by the median of all printed spots for each fluorescent channel separately. Next, the Cy5

and Cy3 signal intensities for each spot were divided by the median value of that spot within the print lot. The threshold value of a specific antibody response for each antigen was defined as four standard deviations above the average of 70 empty phage controls. Statistical significance was calculated using χ^2 test.

The validation of SPAG17-A1 splice variant specific antibodies was done as described in the original paper III. Briefly, an antigen array comprising phage particles in five replicates expressing the dominant transcripts of SPAG17 called SPAG17-A, its minor transcript called SPAG17-A1, an unrelated antigen HORMAD1 and empty phages was tested with serial 3-fold dilutions of two SPAG17-A1 positive gastric cancer sera. The normalised values were further normalized against StrepII signal detected by anti-Sterp II tag antibody (StrepMAB-Immo, IBA, Germany) in order to correct for variations in the copy numbers of recombinant proteins per phage particle.

2.3.2 Characterisation of unnatural ORF antigen antibody specificity

The confirmation of the recognition of the unnatural ORF phage-displayed proteins as compared to their natural ORF counterparts was done as described in original paper I. Briefly, antigens translated both, in their natural ORF as well as in the reading frame displayed on serum-reactive phage clones were cloned in a prokaryotic GST expression vector pGEX- 4T-1 (GE Healthcare, USA). Purified recombinant proteins (500 ng) were used for Western blot analysis and detected by 1:20 000 diluted HRP conjugated goat anti-GST antibody (GE Healthcare, USA) or 1:200 diluted patient's sera followed by incubation with HRP conjugated goat anti-human IgG antibody 1:10000 (SIGMA-Aldrich, Germany) using ECL Plus Western Blotting Detection Reagents (GE Healthcare, USA).

2.4 Extraction of RNA and preparation of cDNA

Bead-based tissue homogenisation was performed by using the FastPrep-24 instrument and Lysing Matrix A (MP Biomedicals, USA) in 1ml of TRIsolution® (Applied Biosystems, USA) followed by the extraction of total RNA according to manufacturers protocol. RNA from melanoma cell lines was kindly provided by Skin Cancer Unit in German Cancer Research Center. RNA of various normal tissues was purchased (Applied Bisystems, USA; Biocat, Germany). cDNA was synthesized by random hexamer priming from 2ug of total RNA by using RevertAidTM First Strand cDNA Synthesis Kit (Fermentas, Lithuania) according to manufacturers protocol.

2.5 Expression analyses

2.5.1 mRNA expression analysis

Qualitative mRNA expression analyses were done as described in original papers II and III. Briefly, RT-PCR reaction mixtures contained 1x reaction buffer, 2.5mM MgCl2, 0.1μM primers, 0.75U Taq DNA polymerase (Fermentas, Lithuania), and $1/60^{th}$ of the prepared cDNA. Amplification of all target and reference genes was performed at the same cycling conditions (30 s at 94°C, 30 s at 60°C, 30 s at 72°C), except for the number of cycles that was adjusted individually according to mRNA abundance. Table 2 displays the used primer sequences.

Quantitative mRNA expression analyses were done as described in original papers II and III. Briefly, real-time RT-PCR (qPCR) reactions were performed using $1/60^{th}$ of cDNA reaction mixtures, ABSolute Blue™ SYBR green Low ROX (Thermo Scientific, USA) and ABI7500 sequence detection system (Applied Biosystems). Appropriate primer concentrations were established by cDNA 4 log serial dilution curves to ensure amplification efficiency over 95%. To normalize the expression data in the sets of various normal or cancer tissues a normalisation factor was calculated for each cDNA from the

expression values of three most stable reference genes in each set determined among 7 most often used house keeping genes (GAPDH, ACTB, POLR2A, TUB3A, TBP, YWHAZ, PGK1) by using geNorm software. The calculated most stable genes were (i) ACTB, POLR2A, TUB3A in the set of various normal tissues, (ii) YWHAZ, PGK1, GAPDH in the breast cancer set, (iii) YWHAZ, ACTB, TBP in the gastric cancer set, and (iv) YWHAZ, TBP, POLR2A in the melanoma set (Figure 4). All reactions were performed in duplicates. The expression level of each target gene was determined relative to its expression in testis. Statistical analysis of expression data was performed by non-parametric Mann-Whitney test.

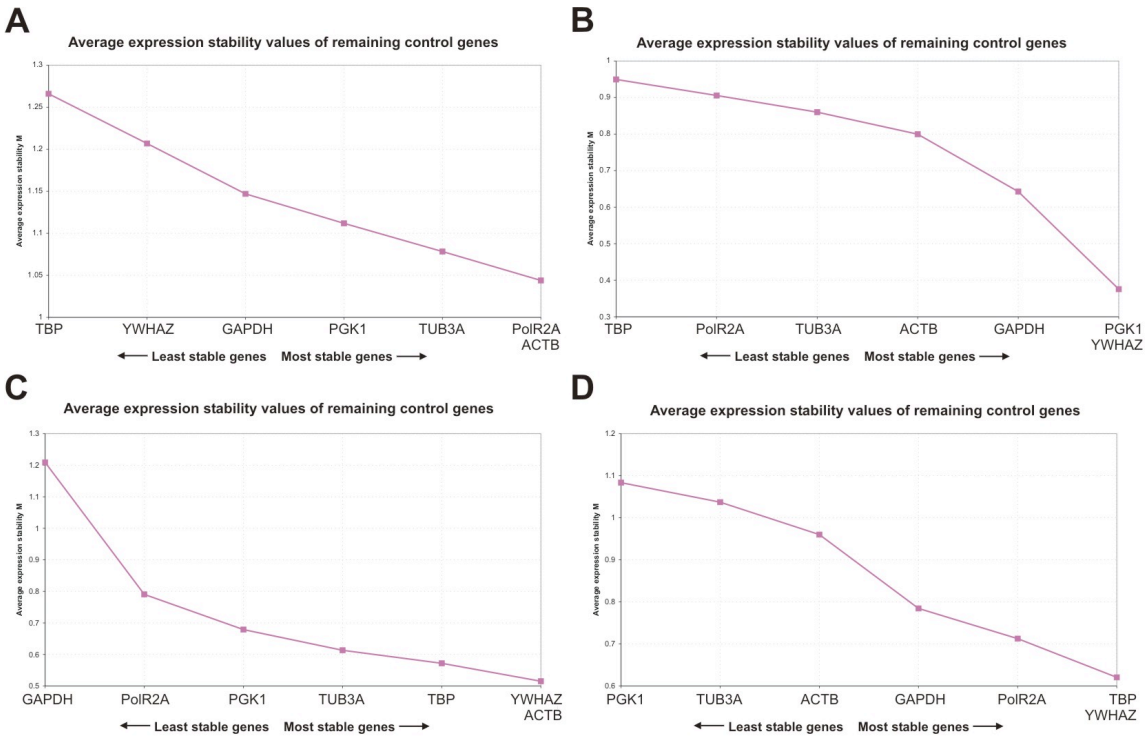


Figure 4. The analysis of most stable reference genes and calculation of normalisation factors in various cDNA sets.

GeNorm software was used to calculate the pair-wise variations of control gene expression among various tissues. Graphs display the stability of expression of seven most often used house keeping genes in sets of various normal tissues (A), breast cancer (B), gastric cancer (C), and melanoma-normal tissue sets (D).

2.5.2 Protein expression analysis

The level of SPAG6 and SPAG8 protein expression was studied by immunohistochemistry (IHC) as described in the original paper III. Briefly, paraffin embedded TMAs of various normal tissues as well as melanoma, lung, gastric and breast cancers (ISU ABXIS, Korea; US Biomax Inc, USA) were processed according to manufacturer's protocol. Commercial antibodies: mouse monoclonal anti-SPAG6 antibody (SantaCruz Inc, USA) and rabbit polyclonal anti-SPAG8 antibody (Proteintech Group, Inc., USA) were used at the dilution 1:50 and incubated overnight at +4°C. Corresponding secondary antibodies – HRP conjugated anti-mouse IgG antibody and anti-rabbit IgG antibody, (SIGMA-Aldrich, Germany) were used at the dilution 1:50 for 1h at +37°C. Followed by dehydration, xylene saturation, mounting in Canada Balsam (SIGMA-Aldrich, Germany) and examination under light microscope (Leica DM3000, Germany).

2.6 Evaluation of spontaneous antigen-specific CTL responses

2.6.1 Peptide predictions

Peptides that could bind MHCI were predicted for tumour antigens LRRC50, SPAG8 and NY-ESO-1 using the open access software SYFPEITHY (http://www.syfpeithi.de/), BIMAS (http://www-bimas.cit.nih.gov/molbio/hla_bind/), and Immune Epitope Database (IEDB) software tools (http://tools.immuneepitope.org/analyse/html/mhc_binding.html) for the most common alleles of the two MHCI genes HLA A and B in the European Caucasians potentially covering 80% of population (A0201 (27%), A03 (>14%), A0101 (16%), B0702 (14%), B08 (12%), B4402 (11%)). Peptides with the highest predictive score for each allele as well as the published MHCI-binding peptides of NY-ESO-1 were purchased (purity >90%) (PiProteomics, USA) (see Table 3 for the obtained peptides). Hydropathicity parameters of peptides were determined using EXPASY Proteomics server software tool ProtParam (http://www.expasy.ch/tools/protparam.html), and only hydrophobic peptides were diluted in DMSO, while hydrophilic peptides were diluted in ice-cold PBS to diminish the possibility of oxidation of M and C residues caused by DMSO. Peptide pools were prepared for each antigen containing 20μM each peptide (20x stock), and immediately divided in aliquots on ice to avoid degradation by repeated freeze/thaw cycles.

For positive control stimulations commercially available viral peptide pool (32 peptides from CMV, EBV) was used (CTL Ltd., USA).

2.6.2 *In vitro* pre-sensitisation

To detect tumour antigen-specific effector T cells of low frequency in PBMCs (<0.01%), it is necessary to amplify the existent effector cells prior to functional analyses, and to do that (i) they need to be separated from the total PBMCs, (ii) antigen-presenting cells need to be prepared, and (iii) target cells for effector function analyses need to be prepared.

For each patient one cryovial of PBMCs (5-10x10^6 cells) was rapidly thawed in 37°C water bath, reconstituted in 10ml of warm complete T cell medium comprised of RPMI1640 medium with 10% human AB serum, 1% non-essential amino acids, 1% L-glutamine, 1% antibiotic/antimycotic, 1% sodium pyruvate, 25U/ml recombinant human IL-2 (all from Invitrogen, Norway), adding the medium slowly drop by drop to avoid osmotic shock of the cells caused by the concentration drop of DMSO. Cells were then centrifuged at 330g for 10 minutes, reconstituted in warm complete medium and counted using EB/AO fluorescent dyes under a UV microscope. CD8+ T cells were extracted from total PBMCs by Dynal® CD8-specific magnetic beads (Invitrogen, Norway) according to the manufacturer's protocol. Briefly, total PBMCs were incubated in 1.5ml tubes with anti-CD8+ antibody coated magnetic beads on ice for 20 minutes with gentle tilting and agitation at 4°C to decrease the phagocytic activity of cells. Cells bound to the beads were separated by placing the tubes in a magnet, aspirating off the non-bound cells, and washed three times while on the magnet. To release cells from the beads they were incubated for 40 minutes at room temperature (RT) with an antibody against the variable region of bead-bound anti-CD8 antibody that outcompetes the attached CD8+ T cells. The released cells were collected by placing the tube in a magnet to fix the released beads and the cells were transferred to a new tube for washing and counting. The same principle was used to separate the CD4+ T cells from the CD8- PBMCs using beads conjugated with anti-CD4 antibody. 4x10^5 CD8+ T cells were transferred to one well of a 96-well plate in 100 μl of 2x complete T cell medium for stimulation with tumour antigen, and 4x10^4 CD8+ T cells were used for stimulation with positive control viral peptides.

CD8-CD4- PBMCs were used as autologous antigen-presenting cells. First, $4x10^5$ cells were reconstituted in 100 µl serum free CTL-Test medium (CTL Ltd., USA) and incubated with the pool of peptides containing 1µM each peptide for 1h. During this time the added peptides bind to the MHCI molecules on the surface of PBMCs as they are in a great molar excess replacing peptides that were bound naturally. Next, the cells were irradiated with 30Gy using linear photon accelerator "Oncor Impression Plus" (Siemens, Germany) at Pauls Stradins Clinical University Hospital to prevent them proliferating in culture, and added to the prepared CD8+ T cell well for stimulation. The CD8+ T cell/APC stimulation culture was grown for 10 days by replacing half of the medium with fresh medium every 2-3 days.

All of the extracted CD4+ T cells were used for the preparation of autologous target cells (PHA blasts). After separation from total PBMCs, they were seeded in complete T cell medium in a well of a 96-well plate, not exceeding $5x10^5$ cells per well. Their proliferation was stimulated with 10 µg/ml (final concentration) phytohemagglutinin PHA-L (SIGMA-Aldrich, Germany), and cells were grown in parallel with the CD8+ T cell culture. Cells were fed as described above and split if necessary.

2.6.3 ELISPOT assay

The presence of antigen-specific effector cells was analysed after 10 days of *in vitro* propagation by IFNγ ELISPOT (Mabtech, Sweden). The principle of this detection system is similar to ELISA. A capture antibody specific to the cytokine of interest is sorbed into the nitrocellulose-covered bottom of the ELISPOT plate wells. If there are activated effector cells in the sample, the produced cytokine will be captured and the presence of activated cytokine-secreting cells can then be visualized by using chromogen like HRP or SAP-conjugated detection antibody against the target cytokine (see Figure 5 for schematic display).

On the day of analysis the complete medium from each CD8+ T cell stimulation well was removed as much as possible not disturbing the cells and reconstituted in 200 µl of serum free CTL-Test medium (CTL Ltd., USA). 20 µl of the virally stimulated cells and 60 µl of cells stimulated with tumour antigen peptides were seeded in three wells of ELISPOT plate for duplicated activation analysis and one negative control. Amplified autologous PHA-stimulated CD4+ T (target) cells were washed and counted and 10^5 cells were pulsed with peptide pool in 100 µl serum free medium as described above and added to each activation analysis well. Autologous target cells without peptide were seeded in each negative control well of the ELISPOT plate. Effector/target cell culture was incubated in the ELISPOT plate overnight avoiding any movement of the plate during this time.

On the next day the cells were removed from ELISPOT plates and the signal detection was performed according to the manufacturer's protocol. Briefly, wells were washed 5 times with PBS, incubated with 1 µg/ml biotin-conjugated anti-IFNγ detection antibody for 2h at RT, washed again 5 times, and incubated with 1:1000 Streptavidin-SAP conjugate for 1h at RT, washed as previously and the presence of IFNγ was visualized by BCIP/NBT-plus substrate solution. To stop colour development, plates were extensively washed in tap water. The coloured spots were analysed after the plates were completely dry. The plates were scanned, quality checked and analysed by the ImmunoSpot® ScAnalysis service (CTL Europe GmbH, Germany). Plates were scanned using CTL-ImmunoSpot® Analyzer and "ImmunoCapture™Software". Counting was made with "ImmunoSpot™ Professional Software" version 5.0.3. in the Smart-Count Mode (all from CTL Ltd, USA).

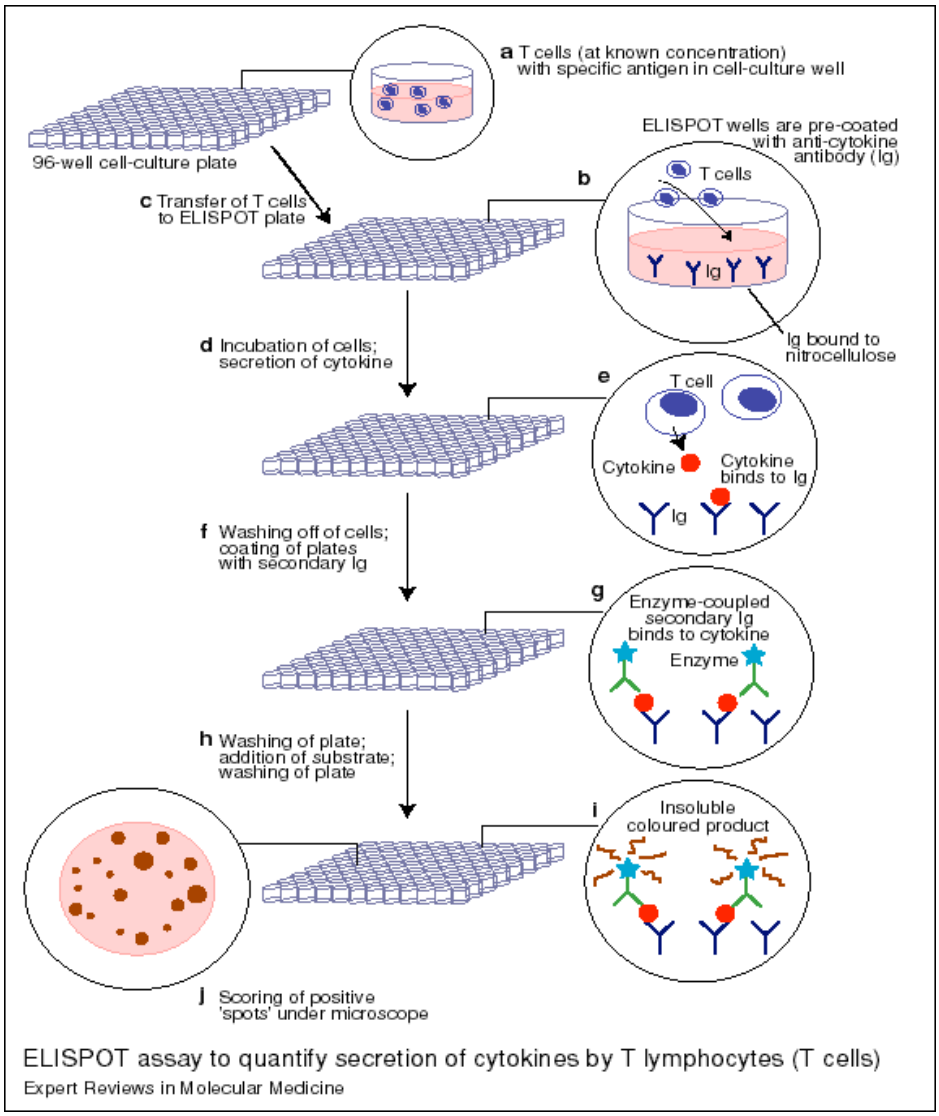


Figure 5. Scheme of ELISPOT analyses .
Adapted from http://www.microvet.arizona.edu/courses/MIC419/ToolBox/elispot.html.

Table 1. Clinical information and collected biological samples of patients used in mRNA expression analyses and T cell activation assays.

Cancer type	Patient number	Type of collected material				Clinical information						
		T	N	S	PBMC	TNM	Tumour stage	Grade	Histological type	Localisation	Age	Gender
Breast	1	+	+	-	-	100	I	medium	NA	1+0*	61	F
	2	+	+	-	-	310	IIIA	high	NA	1+0	73	F
	6	+	+	-	-	210	II	medium	NA	1+0	57	F
	7	+	+	+	-	210	IIB	medium	NA	5	46	F
	8	+	+	+	-	310	III	medium	medullar	4+1	45	F
	9	+	+	+	-	NA	NA	NA	NA	0+4	40	F
	10	+	+	+	-	110	IIA	medium	ductal	4	59	F
	11	+	+	+	-	230	IIIC	medium	ductal	1+0	48	F
	12	+	+	+	-	4xx	NA	NA	ductal	1+4+2	47	F
	13	+	+	+	-	100	I	medium	ductal	4	47	F
	18	+	-	+	+	210	II	low	ductal	2	71	F
	19	+	+	+	+	110	IIA	medium	ductal	3	68	F
	21	+	+	+	+	100	I	high	ductal	4	61	F
	22	+	+	+	+	100	I	medium	ductal	5+3	76	F
	23	+	+	+	+	120	IIIA	medium	lobullar	3	65	F
	24	+	+	+	+	200	II	medium	Mucinous	2+4	78	F
	25	+	+	+	+	220	IIIA	medium	ductal	4	48	F
	26	+	+	+	+	200	II	high	lobullar/papillar	5	64	F
	27	+	+	+	+	400	III3	high	ductal	1	NA	F
	28	+	+	+	+	NA	NA	NA	NA	NA	NA	F
	30	+	+	+	+	NA	NA	NA	NA	NA	NA	F
Colon	100	+	+	+	-	NA	NA	NA	NA	NA	NA	NA
	102	+	+	+	-	300	II (Duke's)	medium	tubular/papillar	rectum	61	M
	103	+	+	+	-	300	II	medium	adenocarcinoma	rectosigmoid	68	M
	21	+	+	+	-	310	IIB	medium	adenocarcinoma	ascendens	72	M
	22	+	+	+	-	300	IIB	medium	adenocarcinoma	rectum	59	F
	27	+	+	+	-	400	IIB	medium	adenocarcinoma	sigmoid	61	F
	30	+	+	-	-	310	IIC	medium	adenocarcinoma	caecum	70	F
	33	+	+	-	-	300	IIB	low	adenocarcinoma	sigmoid	77	F
	425	+	+	+	-	100	1	medium	adenocarcinoma	rectosigmoid	70	M
	437	+	+	-	-	310	IIIA	high	ulcerative adenocarcinoma	sigmoid	29	F
Gastric	201	+	+	+	-	NA	NA	low	NA	antrum	NA	NA
	22	+	+	-	-	301	IV	high	intestinal	antrum	61	F
	27	+	+	+	-	300	II	NA	mucinous	antrum	75	F
	31	+	+	+	-	310	IIIA	high	signet ring	corpus	62	F
	33	+	+	+	-	301	IV	medium	diffuse	pylorus	70	F
	43	+	+	+	-	320	IIIB	high	mucinous	corpus	67	M
	56	+	+	-	-	310	IIIA	high	diffuse/signet ring	pylorus	65	F
	411	+	+	-	-	300	II	high	intestinal	antrum	72	M

Table 1. Continued.

Cancer type	Patient number	Type of collected material				Clinical information						
		T	N	S	PBMC	TNM	Tumour stage	Grade	Histological type	Localisation	Age	Gender
Gastric	436	+	+	-	-	200	IB	low	intestinal	corpus	69	M
	438	+	+	-	-	200	IB	medium	intestinal	corpus	74	M
	440	+	+	-	-	310	IIIA	high	diffuse	corpus	NA	NA
	446	+	+	-	-	310	IIIA	high	NA	corpus	64	M
	449	+	+	-	-	300	II	medium	intestinal	fundus	47	F
	451	+	+	-	-	320	IIIB	high	tubular	corpus/fundus	78	F
	461	+	+	-	-	100	IA	NA	NA	antrum	81	F
	465	+	+	-	-	300	II	high	diffuse	corpus	61	F
	478	+	+	+	+	311	IV	high	diffuse	corpus	58	M
	479	-	-	+	+	321	IV	high	diffuse	antrum	67	F
	480	+	+	+	+	100	IA	low	diffuse	antrum	72	F
	482	+	+	+	+	100	IA	medium/high	intestinal	corpus	62	M
	483	-	-	+	+	301	IV	medium	mucinous	cardia	61	M
	484	+	+	+	+	400	IIIA	NA	NA	corpus	67	F
	485	-	-	+	+	411	IV	high	adenocarcinoma	cardia	65	M
	486	-	-	+	+	410	IV	medium	intestinal	cardia	82	M
	487	-	-	+	+	NA	NA	NA	NA	pylorus	72	M
	488	+	+	+	+	200	IB	high	intestinal	antrum	66	M
	489	+	+	+	+	200	IB	medium	intestinal	corpus	42	M
Lung	13	+	-	-	-	320	IIIA	NA	NSCLC, squamous cell	NA	NA	M
	16	+	-	-	-	410	IIIB	NA	NSCLC, squamous cell	NA	NA	M
	27	+	-	-	-	NA	IIB	NA	NSCLC, squamous cell	NA	NA	M
	28	+	-	-	-	NA	IIB	NA	NSCLC	NA	NA	M
	31	+	-	-	-	NA	IIIA	NA	NSCLC	NA	NA	M
	38	+	-	-	-	NA	IIIA	NA	NSCLC	NA	NA	M
	6	+	-	-	-	320	IIIA	NA	NSCLC, squamous cell	NA	NA	M
Melanoma	1	+	+	+	-	300	II	medium	pigmented epithelial cell	skin	62	F
	2	-	+	+	-	410	V (Clark's)	NA	pigmented epithelial cell	skin	44	F
	4	+	+	+	-	43x	V	NA	non-pigmented nodular	skin	80	F
	5	+	+	+	-	200	IB	NA	pigmented epithelial cell	skin	73	F
	6	-	+	+	-	200	IV	NA	pigmented ulcerative	skin	67	M
	8	+	+	+	-	100	IV	NA	pigmented epithelial/superficial	skin	75	F
	9	+	+	+	-	400	III	medium	pigmented nodular/epithelial, spindle cell	skin	74	F
	273	+	-	+	-	NA	NA	NA	NA	NA	NA	NA
	445	+	-	+	-	NA	NA	NA	NA	NA	NA	NA
	525	+	-	+	-	NA	NA	NA	NA	NA	NA	NA
	1643	+	-	+	-	NA	NA	NA	NA	NA	NA	NA
	1816	+	-	+	-	NA	NA	NA	NA	NA	NA	NA

Table 2. Primers used for the qPCR amplification.

Gene	Forward primer 5'-3'	Reverse primer 5'-3'
PGK1	CTTAAGGTGCTCAACAACATGG	ACAGGCAAGGTAATCTTCACAC
POLR2A	GGGTCATCTTCCCAACTGGAG	CACCAGCTTCTTGCTCAATTCC
TBP	CCACTCACAGACTCTCACAAC	CTGCGGTACAATCCCAGAAC
ACTB	AATCTCATCTTGTTTTCTGCGC	AGTGTGACGTGGACATCCG
GAPDH	GGGTCTTACTCCTTGGAGGC	GTCATCCCTGAGCTAGACGG
TUBA3-R	TATGGCAAGAAGTCCAAGCTG	TACCATGAAGGCACAATCAGAG
YWHAZ	CCTGCATGAAGTCTGTAACTGAG	GACCTACGGGCTCCTACAAC
ACTR2	ATTGCTGGGAGGGATATAAC	TTTCTAAGGCCAGTTTCTGC
AIF1	CATGTCCCTGAAACGAATGC	TAGGATGGCAGATCTCTTG
ALLC	ACTGCTTCCAGTGACCAAGT	GTGTGTTAAGGGTTTGCTTTG
C11ORF20	AGCCTTCGAAAGTGTCGCTC	CCTCGCAGACCTCGGACAG
c16orf82	GAATGAAGCAGGAAGGAGAG	GCTGCACGCTGGAAACAT
C21ORF66	CTGGAGAACAGGATGAAGAG	CTATAAGGAATGCCATAGG
CCDC92	AGCTGGAAGCCCAACTGAAAG	CAGGTAGGCGATGGTGCTG
CFL1	GGTATGCTCTGCAAGGCCTC	CTTCTTGCGCTTCTTCACCTC
Clone #167	TGGCTTTGATGTTGAGTCTCC	AAAATACGAAAACCAGTCAGG
Clone #200	CCAAATAGCTGGATTACAGAA	GTGGTTGATGTCATTGTTTG
Clone #232	GAGGCGAGGGGCAGAATC	GCACTTAATGAACGAAGGGGAG
COPS4	CTGGAGCAGAACAACTAGAAG	CAAATGAAATGACTCTAGGCTG
COX6B2	CCAGAGCCTTGACGAATGCAC	CAGGAGGCTGCTGTGGCTG
ESCO2	CTGCAATCTGCGGGATCAGTC	GTACTGTGTTGCAAACAGC
EVI5L	AGCTACAGGAGCAGCAGG	GCGCTTGTCTCTGGGTTC
LIG2	AGGAGTGGAATGGAGTGGTG	TGGGAGAGGTGTCAGAGAGG
LO392843	CCCCGCTGAGGTCTTTGTG	CCAAATGTATCCCTATGATCTG
LOC284861	TTCTCCAGTGTCAATGGCTCA	ACACTGTACAAGACATCTGTAA
LRRC50	CAAAGAGATGCTGCACCACTC	CTATGATGCTTTCGGTGCTGG
NOL8	TTTACCACAGTCCTCATTCC	ACAAACAGAGGATCAGTAGC
R3HDM2	GAGCAGCACAGACAGCGAAC	GACCTGGTTGCACACTTCTG
RFWD2	AGTCTCTAGTATTGAATTTGACC	CATTCTCAGGGTAATGAATATCC
RNF14	GGAACTCCCATAGAGAAATTAG	CCCAAATATCGTCGTCAACATC
RPL7A	GAATTTTGGCATTGGACAGG	CTGGGTGAACTGGTTAATCG
RPLP1	CCAACGTCAACATTGGGAG	GGTTTAGTCAAAAAGACCAAA
RPS19	GGAAAAGGACCAAGATGGC	GTTCTAATGCTTCTTGTTGGC
SENP1	CAGTGAAGCATTTCGCCTGAC	AACGTTTCACTGCCTGATAACC
SLU7	TCCAGATGAACATGTCCAGCC	TCTCCCCACTGGTGTTTTGG
SPARC	GCAGAGGTGACTGAGGTATC	CCGTGTTTGCAGTGGTGGTT
TEF-ASV	AGGCGCGCCTCGAATCTTC	AGCAAACTTGTGCTTCCGAGG
ZNF282	GAAGAACCTTGTGTGTGGGAG	CAGAAATTGGGGAGTCGGT
SPAG1	GAAAAGCATCTTCAAGCCTTGG	GGAGGTCAAGCACCAAGTTTG
SPAG4	TGGGTCTCCAGTAGTCTCTGA	TCCTCTGCACGACCAGTCG
SPAG5	CGCAGAGCAGGTTCAAACAC	GGAGAGGCACTTGAATGGGA
SPAG6	AGCAATGGCAGTCATCATTTC	GGATGAATGGTCGGGAACTT
SPAG8	CAAGCATGCAGGATGGCTCT	ATGGCTTCACGCTTCCCTCG
SPAG8-e2L	CTACAACTGGGAGGAAGAGAG	GTGGCTGGTACGAGTCTTTC
SPAG9	AGACCCGAGTGGAATCTTTAG	GTTGATCACTCCCTGAGAGC
SPAG9-C	CTCATACCAGCCTGAAGGTC	CCATCGGGTCCTTTGATCTT
SPAG11B	CACCCAGCCTCACTCCATC	CACTTTGCCCTGGAGAATGG
SPAG12	CAAGAAGCTACTGGACCTCG	GATGCACAAACACGCGAGG
SPAG13	GCACCACTGACAATACCATCC	CACGACTATCAACACTGTCACT
SPAG15	GTTGCTCTGGGTGCTTCTG	GGTCCTCGTTCCTCACACA
SPAG16	CGGAAAACAGTTCTTCCTTC	AGACTGAAAGCAATCAAGAG
SPAG16-L	CTGTCTATATGGGATGCAAGAAC	GACCGTTACTCCACTTAAAACTA
SPAG17	AACAGAAATCCTCAAGTGTGC	TGTGTTCACTTTTCCTCCAAC
SPAG17-A	GGAACATTGCTCTCCACTCC	GCTAATCGTCTTCTCCTCGC
SPAG17-A1	CAACATGAGTCTCTGGGTAA	GCTAATCGTCTTCTCCTCGC
SSX2	GCTCAAATACCAGAGAAGATCC	GTGGCCTTGAAACCTAGTTTAG

Table 3. Predicted MHCI biding peptides used for T cell activation assays.

LRRC50	SPAG8	NY-ESO-1
EEINDPKEI	EMGSKKTV	RLLEFYLAM
GPRMTKSSL	GSRSRSRSL	MPFATPMEA
EPLQKLDAL	KLMVCYETL	AADHRQLQL
FPKDRACAE	MLLQHQICK	SLLMWITQC
AEKEERQQW	TPFRKNCSF	APRGPHGGAA
LLLSSPVEV	LSLGKLLPY	APPLPVPGV
SPRPLIQEL	KLLPYEPENY	PLPVPGVLLK
CAEAWARGGY	YQLGEISSL	KEFTVSGNIL
TLSNIFAVSK	QPPGNVYWPL	RGPESRLLEFY
LSDDSDPELDY	AEQEPTRKLF	
	EVESVTHHDY	
	APQLPTWWPL	
	PTQVPAAEDY	
	AEDYLTWKEW	
	RLCHQEMGSK	
	TPVPLSLGKL	

3 Results

The results are presented here as original publications and as unpublished results.

The author's contribution to the enclosed original publications:

Original paper I

Evaluation of T7 and lambda phage display systems for survey of autoantibody profiles in cancer patients. Kalniņa Z, Siliņa K, Meistere I, Zayakin P, Rivosh A, Abols A, Leja M, Minenkova O, Schadendorf D, Linē A. J Immunol Methods. 2008 May 20;334(1-2):37-50. Epub 2008 Feb 21.

Contribution: performed experimental work described in methods chapter 2.4. (Production of GST fusion proteins and Western blot analysis) and used for the result chapters 3.2. (Comparison of serum reactivity against TAA displayed on lambda and T7 phages) and 3.5. (Sequence analysis and characterisation of the identified antigens).

Original paper II

Molecular analysis of serologically defined tumor antigens reveals novel cancer-testis antigens as potential targets for cancer immunotherapy. Siliņa K, Zayakin P, Kalniņa Z, Ivanova L, Meistere I, Stengrēvics A, Wex T, Inderberg Suso E M, Leja M, Malfertheiner P, Gaudernack G, Schadendorf D, Linē A (manuscript submitted to *Cancer Immunology, Immunotherapy*)

Contribution: prepared the tumour tissue RNA and cDNA collection, performed antigen selection and mRNA expression analyses, and partly the serological screening of antigen microarrays; wrote the manuscript.

Original paper III

Sperm associated antigens as targets for cancer immunotherapy: expression pattern and humoral immune response in cancer patients. Siliņa K, Zayakin P, Kalniņa Z, Ivanova L, Meistere I, Endzeliņš E, Ābols A, Stengrevics A, Leja M, Ducena K, Kozirovskis V, Linē A. J Immunother. 2011 Jan;34(1):28-44.

Contribution: prepared the tumour tissue RNA and cDNA collection, performed mRNA and protein expression analyses, growing and purification of recombinant phages, antigen microarray screening with sera from leucocytic leucosis, breast, colon, lung, thyroid cancer and partly melanoma, gastric and prostate cancer patients; wrote the manuscript.

Original paper IV

Alterations of pre-mRNA splicing in cancer. Kalnina Z, Zayakin P, Silina K, Linē A. Genes Chromosomes Cancer. 2005 Apr;42(4):342-57.

Contribution: prepared the graphical information and partially the manuscript.

Original paper V

ELISPOT assays provide reproducible results among different laboratories for T-cell immune monitoring-even in hands of ELISPOT-inexperienced investigators. Zhang W, Caspell R, Karulin AY, Ahmad M, Haicheur N, Abdelsalam A, Johannesen K, Vignard V, Dudzik P, Georgakopoulou K, Mihaylova A, Silina K, Aptsiauri N, Adams V, Lehmann PV, McArdle S. J Immunotoxicol. 2009 Dec;6(4):227-34.

Contribution: performed experimental work as one of the international laboratory participants described in the methods chapter "ELISPOT assays" and used for the result chapters "Establishing Inter-Laboratory Variability of Test results" and "Influence of Serum on the Test Results".

3.1 Identification of cancer antigens and validation of autoantibody responses

ELSEVIER

Journal of Immunological Methods 334 (2008) 37 – 50

JIM Journal of Immunological Methods

www.elsevier.com/locate/jim

Research paper

Evaluation of T7 and lambda phage display systems for survey of autoantibody profiles in cancer patients

Zane Kalniņa [a], Karīna Siliņa [a], Irēna Meistere [a], Pawel Zayakin [a], Alexander Rivosh [a], Artūrs Ābols [a], Mārcis Leja [b], Olga Minenkova [c], Dirk Schadendorf [d], Aija Linē [a,*]

[a] *Biomedical Research and Study Centre of Latvia, Riga, Latvia*
[b] *Faculty of Medicine, University of Latvia, Riga, Latvia*
[c] *Kenton Labs, c/o Sigma-Tau, Pomezia, Italy*
[d] *German Cancer Research Center, Skin Cancer Unit, Heidelberg, Germany*

Received 3 January 2008; received in revised form 28 January 2008; accepted 29 January 2008
Available online 21 February 2008

Abstract

In the current study we attempted to evaluate the suitability of T7 Select 10-3b and λKM8 phage display systems for the identification of antigens eliciting B cell responses in cancer patients and the production of phage-displayed antigen microarrays that could be exploited for the monitoring of autoantibody profiles. Members of 15 tumour-associated antigen (TAA) families were cloned into both phage display vectors and the TAA mini-libraries were immunoscreened with 22 melanoma patients' sera resulting in the detection of reactivity against members of 5 antigen families in both systems, yet with variable sensitivity. T7 phage display system showed greater sensitivity for the detection of antibodies against members of CTAG, MAGEA and GAGE families, both systems showed equal performance in detecting the reactivity against MAGEC and SSX2 while only λKM8 allowed the detection of anti-CTAGE5 antibodies. The biological properties of both phages turned out to be equally suitable for the production of antigen microarrays however in line with the plaque assay the sensitivity for the detection of various autoantibodies differed between the vectors. However, presumably due to the higher variability of the background signals in the microarray assay, it turned out to have comparable, in some cases even slightly lower sensitivity than the plaque assay.

Next, we explored the repertoire of antigens that could be identified by screening T7 phage-displayed testis cDNA library with sera from melanoma patients. From the 243 antigens identified, only 24 represented known genes translated in their natural reading frame and included known TAAs like Annexin XI-A and a novel potential CT antigen SPAG8. Another 12 were uncharacterised genes but the remaining clones contained DNA fragments in non-natural reading frames that most likely represent mimotopes, nevertheless, they may turn out to be valid biomarkers.

Keywords: Autoantibodies; Phage display; Antigen microarray; SEREX; Melanoma antigens

1. Introduction

Circulating autoantibodies against tumour-derived proteins have been observed in the most if not all cancer patients hence they may serve as biomarkers for the screening, diagnosis, prognosis or monitoring of cancer. In fact, autoantibodies against a number of CT antigens, including

Abbreviations: SEREX, serological identification of antigens by recombinant expression cloning; TAA, tumour-associated antigens; pfu, plaque forming units; CT, cancer-testis; UTR, untranslated region, ORF, open reading frame.

* Corresponding author. Biomedical Research and Study Centre of Latvia, Ratsupites Str 1, LV-1067, Riga, Latvia. Tel.: +371 7808208, fax: +371 7442407.
E-mail address: aija@biomed.lu.lv (A. Linē).

0022-1759/$ - see front matter © 2008 Elsevier B.V. All rights reserved.
doi:10.1016/j.jim.2008.01.022

NY-ESO-1, SSX2, MAGEA1 and 3 (Stockert et al., 1998) and several members of CTAGE family (Usener et al., 2003), and melanocyte differentiation antigens such as RAB38 (Zippelius et al., 2007) have been detected exclusively in the sera from cancer patients, thus suggesting they may serve as highly specific diagnostic markers. Autoantibodies against another set of antigens such as GLEA2, PHF3 and endostatin have been shown to correlate with the clinical outcome in glioblastoma and breast cancer patients, respectively (Pallasch et al., 2005; Bachelot et al., 2006). Moreover, appearance of anti-p53 autoantibodies may predict subsequent development of cancer (Li et al., 2005), while their prognostic significance still remains a subject of debate (Soussi, 2000). However, so far the clinical utility of tumour-associated autoantibodies has been hampered by the low frequencies to each particular antigen and the heterogeneity of antibody repertoires in cancer patients. This could be overcome by developing high-throughput techniques that would allow to identify a comprehensive set of immunogenic proteins in a given type of cancer and subsequently analyse the occurrence of antibodies against them in large sets of sera from patients with cancer, autoimmune disorders and healthy donors.

SEREX (serological identification of tumour antigens by recombinant expression cloning) is a widely used technique for the identification of immunogenic proteins in cancer patients that is based on the construction of cDNA expression libraries from tumour tissues, expression of the recombinant proteins in *E. coli* and immunoscreening of the libraries with sera from cancer patients (Sahin et al., 1995). The application of this technique to a variety of tumour entities has led to the identification of more than 2000 genes encoding potential tumour antigens, most of which are deposited in the Cancer Immunome database (http://www2.licr.org/CancerImmunomeDB). However, this approach is extremely time-consuming and labour-intensive and therefore is generally not applicable for the profiling autoantibody responses in multiple serum samples. Recently, several laboratories have exploited various phage display strategies, including M13 filamentous phage (Sioud and Hansen, 2001; Somers et al., 2002), pJuFo system (Fossa et al., 2004), T7 phage (Hansen et al., 2001; Zhong et al., 2004) and lambda phage (Minenkova et al., 2003; Pavoni et al., 2004) for the construction of cDNA expression libraries and successfully applied them for the isolation of cDNAs encoding tumour-associated antigens. A feature that makes the phage display-based strategies particularly attractive for the searching of antigens is that tumour-derived cDNAs are expressed as fusion proteins with one of the phage coat proteins (or linked via a Jun–Fos interaction in case of pJuFo system) and exposed on the surface of the phage thus allowing the selection of serum-reactive phage clones by biopanning, which is much faster as well as more cost and labour-effective approach than the conventional immunoscreening. Moreover, the recombinant phage particles can directly be printed onto glass slides to produce antigen microarrays that allow monitoring the antibody responses against hundreds to thousands of antigens simultaneously using a microlitre of serum (Fernandez-Madrid et al., 1999; Cekaite et al., 2004; Zhong et al., 2005; Wang et al., 2005; Chatterjee et al., 2006). However, the main drawback of the phage display-based systems for the expression of cDNA libraries is the bias in the repertoire of proteins that can be displayed on the surface of the phage resulting from the biological peculiarities of a given phage species. For example, the assembly of M13 phage takes place in the periplasm of *E. coli*, therefore only those fusion proteins that can be exported through the bacterial inner membrane will be displayed on the phage capsid (Hufton et al., 1999). Capsids of lytic phages such as T7 and lambda phages are assembled in the cytoplasm therefore the repertoires of cDNA libraries do not have the biological constraints similar to M13 phage (Krumpe et al., 2006). The capability of a protein to be displayed on the surface of a lytic phage depends mostly on its impact on the assembly process, where the length, biochemical properties, hydrophilicity and hydrophobicity as well as folding characteristics play the major role. Moreover, the copy number of recombinant proteins per phage particle may vary in these systems allowing the assembly of chimerical capsids, thus affecting the sensitivity of the assay.

In the current study we assessed the capability of T7 Select 10-3b and λKM8 phages to display a range of clinically relevant tumour antigens by cloning the members of 15 TAA families into these display vectors, compared the sensitivity of the systems for the detection of autoantibodies by immunoscreening of the TAA mini-libraries with sera from 22 melanoma patients, and evaluated the suitability of these phages for the production of antigen microarrays. Moreover, in order to characterise the repertoire of melanoma associated antigens that can be identified using T7 Select 10-3b system, we constructed T7 phage-displayed testis cDNA library and screened it with sera from 9 melanoma patients resulting in the identification of 243 different sero-reactive phage clones.

2. Materials and methods

2.1. Serum samples

Serum samples from melanoma patients (stage II–IV) were collected at the Skin Cancer Unit, German Cancer Research Center and Latvian Oncology Center after the

patients' informed consent was obtained in accordance with the regulations of local ethics committees. The samples were aliquoted and stored at −70 °C.

2.2. Construction of phage-displayed cDNA expression libraries

2.2.1. T7 and lambda phage-displayed tumour-associated antigen (TAA) mini-libraries

Either the entire ORFs or the antigenic regions predicted using an algorithm developed by Welling et al. (1985) of 15 tumour antigens or antigen families were amplified by PCR using High Fidelity PCR enzyme mix (Fermentas) and cDNA from testis or melanoma tissue as a template. The antigens and the regions chosen for cloning are listed in Table 1, and the sequences of PCR primers are available on request. All reverse primers contained *NotI* site and forward primers contained either *SpeI* or *SalI* sites for cloning into λKM8 or T7 Select 10-3b vectors, respectively. For the amplification of genes from multigene families degenerated primers were designed to amplify as many as possible members of the same gene family. If more than one region of a gene was amplified, PCR products were combined prior to ligation in the vector. PCR products were digested with *NotI* and *SpeI* or *SalI* (Fermentas) and purified using GFX PCR DNA and Gel Band Purification kit (Amersham Biosciences). λKM8 (Pavoni et al., 2004) vector DNA was digested with *SpeI* and *NotI* but T7 Select 10-3b vector (Novagen) was digested with *SalI* and *NotI*, treated with shrimp alkaline phosphatase (Fermentas) followed by phenol/chloroform extraction and ethanol precipitation. Each digested PCR product (5 ng) was ligated into λKM8 or T7 Select 10-3b vector (100 ng), and 1/5 of each ligation mixture was subjected to *in vitro* packaging using 5 µl of Gigapack III Gold Packaging extract (Stratagene) or T7 Select Packaging extract (Novagen), respectively, resulting in mini-libraries of $1–5\times10^5$ pfu for lambda phage and $0.5–5\times10^4$ pfu for T7 phage. The obtained TAA mini-libraries were amplified once using *E. coli* BB4 cells or IPTG-induced BLT5615 cells, respectively. The lysates were centrifuged to remove cell debris and stored at +4 °C or as glycerol stocks at −80 °C.

2.2.2. T7 phage-displayed testis cDNA expression library

Testis cDNA library was constructed using T7 Select 10-3b vector and OrientExpress cDNA library construction system (Novagen) according to the manufacturer's instructions. Briefly, mRNA was isolated from 150 µg of testis total RNA (Ambion) using Dynabeads mRNA purification kit (Invitrogen) and converted to cDNA using HindIII Random primers (5′-TTNNNNNN-3′). Then cDNA was ligated to directional EcoRI and HindIII linkers, digested with the corresponding restriction enzymes and ligated into pre-digested T7 Select 10-3b vector followed by *in vitro* packaging resulting in a

Table 1
Antigens comprising TAA mini-libraries

Antigen family	Number of mRNAs[a]	Regions cloned (nt positions/NCBI RefSeq No)	Different genes found by sequencing 5 random clones	
			T7 Select 10-3B	λKM8
CTAG1B (NY-ESO)	2	86–400; 338–628/NM_001327	CTAG1B, CTAG2	CTAG1B, CTAG2
MAGEA	10	209–618; 590–1137/NM_005362	MAGEA1, 2, 3, 4, 9	MAGEA1, 3, 4, 6
SSX	10	58–401; 58–617/NM_005636	SSX2	SSX1, 2
BAGE	1	195–330/NM_001187	BAGE	BAGE
GAGE	8	84–410/NM_001472	GAGE2, 7, 8	GAGE1, 8
MAGEB	4	98–516/NM_002364	MAGEB2,4	MAGEB2
MAGEC	2	931–1453/NM_016249	MAGEC1, 2	MAGEC1, 2
SPANX	6	57–347/NM_022661	SPANXA2, SPANXE	SPANXD
LDHC	1	245–785/NM_017448	LDHC	LDHC
CT45	4	246–798/NM_001017417	CT45-1	CT45-1
THEG	1	778–1176/NM_016585	THEG	THEG
CTAGE	6	130–1070; 311–1070; 1052–1616; 1589–2354/203354; 11–235/NM_022663	CTAGE1, 3, 5, 5Δex7	CTAGE1, 5
MTA1	1	1394–1772/NM_004689	MTA1	MTA1
TYR	1	494–710/NM_000372	TYR	TYR
MLANA	1	54–408/NM_005511	MLANA	MLANA

[a] Number of known members of the antigen family that theoretically could be amplified with the selected primers.

library of 8×10^6 pfu. The library was amplified once in IPTG-induced BLT5615 cells.

2.3. Selection of serum-reactive phage clones

2.3.1. Immunoscreening of phage-displayed TAA mini-libraries

For screening of T7 phage-displayed TAA mini-libraries, ~10^3 pfu from each library was spotted on gridded LB/carbenicillin agar plates pre-coated with LB top agarose containing BLT5615 cells grown in LB supplemented with 1×M9 salts, 0.4% glucose, 1 mM $MgSO_4$ and carbenicillin (50 μg/ml) to OD_{600}=0.5 and induced with IPTG for 30 min. After ~2 h incubation at 37 °C when the plaques reached ~1 mm in diameter, plates were overlaid with Protan nitrocellulose (NC) filters (Schleicher & Schuell) and incubated for 1 h at 37 °C. For screening of lambda phage-displayed TAA mini-libraries, BB4 cells grown in LB supplemented with 0.2% maltose and 10 mM $MgSO_4$ were plated on NZY agar plates, infected with ~10^3 pfu from each mini-library and the plates were incubated at 37 °C until visible plaques appeared (6–8 h). The filters were blocked with 5% (w/vol) milk powder in TBS, 0.05% Tween 20 for 1 h, and then incubated overnight with 1/200 diluted patients' serum preabsorbed with *E.coli*-phage lysate. The serum-reactive clones were detected by incubating the filters with alkaline phosphatase conjugated anti-human IgG, Fcγ specific secondary antibody (Pierce) and NBT/BCIP (Fermentas). Ten serum-reactive clones from each mini-library were subjected to secondary screenings and subcloned to monoclonality.

2.3.2. Biopanning of T7 phage-displayed testis cDNA library

Approximately 5×10^{10} pfu from T7 phage-displayed testis cDNA library were incubated overnight with 2 μl of patient's serum that had been preabsorbed with BLT5615 and T7 phage lysate coupled with BrCN-activated sepharose. A hundred microlitre of Protein G coated magnetic beads (Pierce) were washed twice with blocking solution (5% milk powder in TBS, 0.05% Tween 20), added to the phage–serum mixture and incubated for 2 h at RT under agitation. The beads were washed 10 times with 1 ml TBS, 0.05% Tween and the bound phages were either amplified and subjected to the second round of biopanning or titrated and used for immunoscreening as described above.

2.3.3. Identification of cDNAs

The inserts of serum-reactive phages were amplified by 35-cycle PCR using primers flanking the insert (T7-Up2: 5′-CTTCGCCCAGAAGCTGCA-3′, T7-Down: 5′-AACCCCTCAAGACCCGTTTA-3′, λKM8-Up: 5′-CAATCTGTGTGGGCACTCG-3′, λKM8-Down: 5′-CGGCTGGTAATGGGTAAAGG-3′) and 1 μl of phage solution as a template. PCR products were purified and directly sequenced using ABI Prism BigDye Terminator v3.1 cycle sequencing kit and 3130 Genetic Analyser (Applied Biosystems). DNA sequences were analysed using BLAST tool at www.ncbi.nlm.nih.gov, Translate tool at www.expasy.org and compared against sequences available at Cancer Immunome Database (www2.licr.org/CancerImmunomeDB).

2.4. Production of GST fusion proteins and Western blot analysis

To confirm the recognition of the phage-displayed proteins, GST fusion proteins were produced and used for standard Western blot analysis. cDNA inserts of serum-reactive clones were amplified by PCR and cloned into a prokaryotic GST expression vector pGEX-4T-1 (Amersham Biosciences) to produce either the natural products of these genes or the out-of-frame peptides displayed by the serum-reactive clones.

Purified recombinant proteins (500 ng) were resuspended in Laemmli sample buffer, denatured for 5 min at 100 °C and separated by SDS/polyacrilamide gel electrophoresis (PAGE) on two gels simultaneously and transferred onto Protan NC membranes (Schleicher & Schuell). The filters were blocked for 1 h in 5% milk powder in TBS, 0.05% Tween 20 and incubated with 1/20 000 diluted HRP conjugated goat anti-GST antibody (Amersham Biosciences) or 1/200 diluted patient's sera followed by incubation with HRP conjugated goat anti-human IgG antibody 1/10 000 (Sigma) and detected using ECL Plus Western Blotting Detection Reagents (Amersham Biosciences).

To determine the copy number of fusion proteins per T7 phage particle, phages were amplified in *E. coli* BLT5615 cells, the phage particles were precipitated from the lysates with 3.5% PEG 8000, 0.4 M NaCl, titrated and ~10^9 pfu of each phage were used for Western blot analysis with 1/5000 diluted rabbit HRP conjugated anti-T7 tag antibody (ABcam) recognising the N-terminus of T7 coat protein 10B. To ascertain that unincorporated recombinant proteins are not co-precipitated with the phage particles, NaCl/PEG precipitated phage solution containing ~10^{10} pfu was filtrated through 100 000 NMWL filter units (Millipore) and the flow-through fraction was run in parallel with the phage samples as a negative control. For quantity calibration a standard curve was constructed from 5 serial 3-fold dilutions of one phage clone. The images were analysed using GelWorks software (Ultra-Violet Products).

2.5. Production of antigen microarrays

Nine corresponding T7 and lambda phage clones isolated by immunoscreening from TAA mini-libraries and four non-recombinant phage clones were grown in BLT5615 and BB4 cells, respectively, until complete lysis. The lysates were clarified by centrifugation and arrayed in 5 replicates onto nitrocellulose-coated 16-pad FAST slides (Whatman) using a QArray Mini microarrayer (Genetix) to generate T7 and lambda phage-displayed antigen microarrays. The microarrays were blocked in 5% (w/vol) milk powder in TBS, 0.05% Tween 20 for 1 h, and then incubated with 1/200 diluted patients' sera (preabsorbed with *E.coli*-phage lysates) for 2 h at room temperature. The slides were rinsed with TBS and washed 4 times in TBS, 0.5% Tween 20 for 15 min each and then incubated with anti-T7 tail fiber (Novagen) or anti-gpV (λ tail protein) monoclonal antibody (kindly provided by Dr. Mauruzio Cianfriglia) at a dilution of 1/10 000 or 1/1500, respectively to determine the amount of phages in each spot. After 3 washes in TBS, 0.5% Tween 20 for 10 min, the microarrays were incubated with Cy5 labelled goat anti-human IgG antibody (1/1500) and Cy3 labelled goat anti-mouse IgG antibody (1/3000) (Jackson ImmunoResearch) for 1 h, then washed thrice in TBS, 0.5% Tween 20, rinsed with distilled water and dried by centrifugation. The microarrays were read using AQuire scanner (Genetix) and the images were analysed using Genetix QScan software. For each spot the mean Cy5 and Cy3 signals were background subtracted, averaged between replicates, and the Cy5/Cy3 ratios were calculated for each antigen and normalised by that of non-recombinant phages. A cut-off value for defining serum-reactive antigens was set as >3 SDs above the mean ratio for all the spots.

3. Results

3.1. Construction of lambda and T7 phage-displayed TAA mini-libraries

Members of 12 CT antigen families and 3 other tumour antigens were cloned into lambda KM8 and T7 Select 10-3b phage display vectors to produce TAA mini-libraries. In λKM8 vector cDNAs are fused to the N-terminus of the coat protein gpD and are separated by a flexible GS linker while in T7 phage they are fused to the C-terminus of the coat protein 10B. The antigen families and individual genes as well as the protein regions chosen for cloning are listed in Table 1. Complete ORFs were amplified for shorter transcripts, while for genes with long ORFs the potential antigenic regions were predicted using an algorithm developed by Welling et al. (1985) and chosen for cloning. In the case of tyrosinase the immunodominant region that has been shown to react with sera from melanoma patients

Table 2
Antigens identified by immunoscreening lambda and T7 phage-displayed TAA mini-libraries and individual sera recognising the respective antigens

TAA mini-library	Antigen	Reactive sera[a]	
		T7 Select 10-3B	λKM8
CTAG	CTAG1B	MA001643, MA000703, MA000161, MA000445, MA00SK, MA00AM, MA00550, MA000513	MA001643, MA000703, MA000161, MA000445
	CTAG2	MA001643, MA000703, MA000161, MA000445, MA00SK, MA00AM, MA00550	MA001643, MA000703, MA000445
	CTAG1B-ORF2	UKRV-Mel31	–
MAGEA	MAGEA1	MA000513	MA000513
	MAGEA2	MA000513	MA000513
	MAGEA3	MA000513	MA000513
	MAGEA12	MA000513	–
	MAGEA2 antisense	MA00WF	–
MAGEC	MAGEC1	MA000703	MA000703
	MAGEC2	MA000703	–
SSX	SSX2	MA000951	MA000951
CTAGE	CTAGE5		MA000273, MA000525
	CTAGE pseudogene	MA000525, MA000273, MA001816, MA001111	–
	CTAGE1 antisense	MA001404	–
GAGE	GAGE3-7 subgroup	MA000445, MA001111, MA001404	MA000445
	GAGE1, 8	MA000445	MA000445
	GAGE2-ORF2	MA00GG	–
	GAGE7-ORF2	MA001643	–

[a] Sera are ranked by the signal intensities in immunoscreening.

but not with sera from healthy donors (Lucchese et al., 2005) was chosen. For the amplification of genes from multigene families, degenerated PCR primers capable of amplifying as many members of these families as possible were designed. The cloning of the respective PCR products into λKM8 and T7 Select 10-3b vectors resulted in TAA mini-libraries in the size of $1–5\times10^5$ pfu and $0.5–5\times10^4$ pfu, respectively. The complexity of the mini-libraries and the insertion of cDNAs in the correct reading frame relative to the phage coat proteins were determined by sequencing 5 random clones from each mini-library (Table 1).

3.2. Comparison of serum reactivity against TAA displayed on lambda and T7 phages

Approximately 2×10^3 pfu from each of the 15 lambdas and T7 phage-displayed TAA mini-libraries were immunoscreened with sera from 22 melanoma patients. Ten sero-reactive phage clones from each sero-reactive TAA mini-library were purified and identified by sequencing their cDNA inserts. Generally, relatively good concordance in the recognition of antigens displayed on lambda and T7 phages was observed — the reactivity against antigens representing CTAG (NY-ESO-1), MAGEA, MAGEC, SSX and GAGE families was detected in both expression systems (yet with different sensitivity) whereas no reactivity against BAGE, MAGEB, SPANX, LDHC, CT45, THEG, MTA1, TYR and MLANA was observed in any of them (Table 2).

In both systems, phage capsids are composed of wild-type capsid proteins and hybrid proteins, however the display density and its regulation are different. T7 Select 10-3b system has been shown to display 5–15 copies of recombinant protein per phage but the rest of the wild-type 10A capsid protein (415 in total) is provided by a plasmid in a complementing host (BLT5615) (Rosenberg et al., 1996). In the λKM8 system, wild-type copies of protein D are provided by a copy of the respective gene in the genome of the phage and the capsid was shown to be composed of ~50% recombinant proteins at least in the case of scFv antibody display (Vaccaro et al., 2006). Since it has been suggested that the display density on T7 phage is variable and depends on the size of the hybrid coat protein (Zucconi et al., 2001), we tried to determine the copy number of fusion proteins per phage particle in order to verify that the lack of reactivity against the mentioned antigens is not due to the failure to display the respective antigens. Individual T7 phage clones encoding 11 different antigens were amplified, the phage particles were precipitated from the bacterial lysates with PEG/NaCl and subjected to Western blot analysis using antibody against the N-terminus of T7 coat protein 10B that should recognise both, the fusion protein and the non-recombinant coat protein. Five 3-fold dilutions of the clone encoding MLANA were used to construct a standard curve and the quantity of each band was calculated using GelWorks software. The fusion proteins of the corresponding length were detected in all phage clones analysed, however the copy number per phage varied markedly being the highest in phages expressing N-terminus of CTAG1B (~17%) and lowest in phages expressing N-terminus of CTAGE5 (~1%) (Fig. 1).

Concerning the sensitivity, T7 Select phage display system showed a greater sensitivity in detecting the reactivity against members of CTAG, MAGEA and GAGE

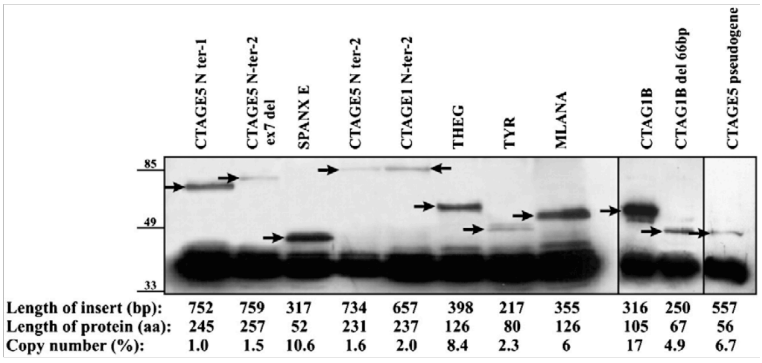


Fig. 1. Copy number of TAA fusion proteins per T7 phage particle determined by Western blotting using antibody against N-terminus of the coat protein 10A/10B. The lower band represents *wt* protein 10A, arrows indicate the fusion proteins of the respective sizes. Quantification of *wt*/fusion protein ratio was done by construction of a standard curve from a five 3-fold dilutions of one phage on separate gels (not shown) and the intensities of bands were calculated using GelWorks software.

antigen families both in terms of signal intensities and the number of positive sera. Both systems showed equal performance in detection of reactivity against MAGEC and SSX2 while only λKM8 allowed the detection of anti-CTAGE5 antibodies.

Eight sera reacted with CTAG1B (7 with CTAG2) when these antigens were displayed on T7 phage but only 4 of them reacted with CTAG1B (3 with CTAG2) when expressed by λKM8. Since the signal intensities for those four sera that did not recognise λ phage-displayed antigens were low, this inconsistency most likely resulted from more efficient display of this particular antigen on T7 phage (in fact, display density of CTAG1B was the highest among all the phages analysed), and not from the inability of lambda phage to display CTAG antigens.

An opposite situation was observed in the case of CTAGE5, where 2 sera recognised its N-terminal fragment when it was displayed on lambda phage but did not react with the corresponding T7 phage. At the same time, 4 sera reacted with T7 phage clones expressing 56 aa polypeptide derived from CTAGE pseudogene containing a stop codon. Western blot analysis showed that T7 phages display ~30 copies of the polypeptide encoded by the pseudogene but only 4 copies of CTAGE5 hence suggesting that very low display density did not allow to detect the presence of serum antibodies against CTAGE5. Alternatively, there could be conformational differences of this antigen when displayed on T7 resulting in poor accessibility of the epitope.

Moreover, 5 sera reacted with T7 phage clones expressing CTAG1B, GAGE2 and GAGE7 ORF 2 peptides and MAGEA2 and CTAGE1 antisense peptides. All these cDNAs contain stop codons when translated as fusion proteins with the coat protein 10B. Since cDNAs are expressed as N-terminal fusion proteins to gpD in λKM8 system, such peptides per definition cannot be displayed on lambda phage. UKRV-Mel-31 serum reacted with 3 different CTAG1B-encoding phage clones containing frame-shifting deletions (most likely resulting from PCR errors) resulting in the expression of CTAG1B-ORF2 peptide, which is well known CD4+ and CD8+ T cell antigen whose immunogenicity is supposed to be associated with the translation of CTAG1B alternative ORF in cancer cells (Slager et al., 2003). Whether or not MAGEA2, CTAGE1, GAGE2 and 7 are also translated in alternative ORFs in cancers and therefore could be considered as tumour specific antigens remains to be determined. The variable display density and the detection of reactivity against the clones expressing out-of-frame peptides could result in false negative and positive calls, respectively, that provides a disadvantage of T7 phage display system in studies aiming to determine the presence of antibodies against definite antigens, therefore λKM8 would be the expression system of choice for these kinds of studies. At the same time, the out-of-frame peptides might turn out to be novel tumour specific antigens or valid biomarkers.

3.3. Suitability of T7 and lambda phages for the production of antigen microarrays

In order to assess the suitability of both phage display systems for the production of antigen microarrays, nine

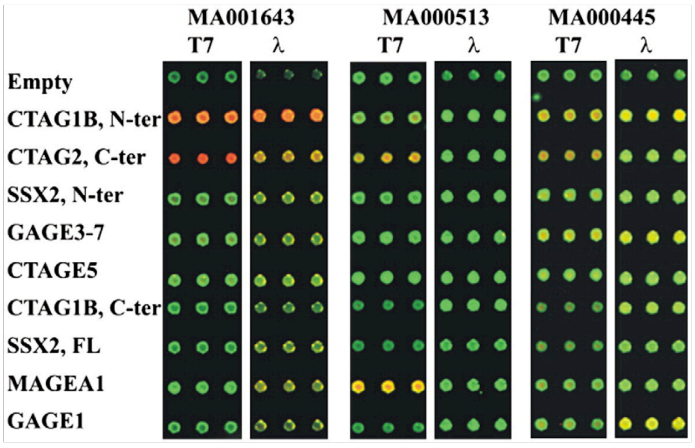


Fig. 2. T7 and lambda phage-displayed antigen microarrays. Nine recombinant T7 and lambda phage clones expressing corresponding TAAs and four non-recombinant phages were amplified, spotted on FAST slides in quintuplicate and tested with sera from melanoma patients (detected with Cy5 labelled secondary antibody) and with monoclonal antibody against tail protein of T7 or lambda phage (detected with Cy3 labelled secondary antibody) to quantify the amount of phages in each spot. Only partial array images are shown.

Table 3
Detection of autoantibodies against T7 and lambda phage-displayed TAAs using microarrays

	MA001643		MA000513		MA000445		MA000951		MA000161		MA000273		MA000525	
	T7	λ	T7	λ	T7	λ	T7	Λ	T7	λ	T7	λ	T7	λ
CTAG1B, N-ter	**37.9**	**17.5**	**28.8**	**3.9**	**8.2**	**5.1**	**2.1**	1.4	**39.3**	**2.8**	0.9	1.5	1.9	0.9
CTAG2, C-ter	**58.8**	3.2	**64.4**	1.5	**9.8**	2.1	1.3	1.0	**43.2**	1.6	1.4	1.6	1.4	0.9
SSX2, N-ter	2.3	1.2	2.9	1.0	2.5	1.7	**3.5**	**1.9**	3.5	1.1	1.1	1.6	1.6	0.9
GAGE3-7	2.2	1.2	3.0	1.0	**5.2**	**4.8**	1.5	1.4	3.4	1.4	1.2	1.6	1.4	1.0
CTAGE5	2.0	1.2	2.5	2.1	1.7	1.8	1.2	1.1	3.3	1.4	0.9	**2.5**	1.3	1.2
CTAG1B, C-ter	4.6	2.7	3.0	0.9	2.3	2.8	1.6	1.2	1.1	1.6	0.9	1.6	1.5	0.9
SSX2, FL	2.3	1.0	2.8	0.6	2.4	1.2	**4.3**	**2.1**	1.8	0.8	1.0	1.4	1.2	0.9
MAGEA1	2.3	1.2	**119.7**	**2.7**	1.7	1.7	1.6	1.4	1.6	1.3	1.2	1.7	1.4	1.0
GAGE1	2.2	1.3	2.8	1.3	3.0	**5.7**	1.6	1.2	2.9	1.7	0.8	1.5	1.5	1.1

Cy5/Cy3 ratios for each antigen were normalised by that of non-recombinant phages and the cut-off value for serum-positive clones was set as >3 SDs above the average of all the spots (marked with bold).

T7 and lambda phage clones encoding corresponding TAAs and 4 non-recombinant phages were amplified and spotted on FAST slides, and the microarrays were tested with 7 serum samples used for the immunoscreening of the TAA mini-libraries (Fig. 2). The variation between replicates was less than 10% for both display systems. Similarly to plaque immunoscreening, T7 Select system showed a higher sensitivity for the detection of autoantibodies against CTAG1B and MAGEA1, both systems detected anti-SSX2 and GAGE antibodies with comparable signal-to-noise ratios but anti-CTAGE5 antibodies were detected only in λKM8 system (yet only in one out of two CTAGE-positive serum samples) (Table 3). In most of the cases, the antigens that were found to be serum-positive by the immunoscreening of TAA mini-libraries were defined as positive using the selected cut-off value (>3 SDs above the average of all spots in the array) in microarray screening. However, although T7 phage clones expressing CTAG1B C-terminus and GAGE1 were identified by plaque immunoscreening using MA001643 and MA000445 sera, respectively, the signal intensities for these clones using the respective sera did not reach the defined cut-off value. Similarly, lambda phage clones expressing CTAG1B and CTAG2 C-termini and CTAGE5 also were called sero-negative against the sera that were used to isolate these clones from TAA mini-libraries. Nonetheless, lowering the cut-off value most likely would result in false-positive calls due to the variability in the background signal intensities of serum non-reactive clones. At the same time MA000951 serum was defined as CTAG1B positive by microarray screening while no reactivity against CTAG1B was detected by plaque assay. Hence the microarray screening and plaque immunoscreening have comparable yet not identical sensitivity for the detection of autoantibodies.

3.4. Selection of sero-reactive clones from T7 phage-displayed testis cDNA library

As the ultimate goal of our study is to identify a comprehensive set of antigens that could be further used for profiling autoantibody repertoires in patients' sera, we next decided to explore the repertoire of antigens that could be identified using T7 Select phage display system. Since the germ cell transcriptome shares many characteristics with cancers, we used testis as a source of RNA for the construction of T7 phage-displayed cDNA expression library and screened it with sera from 9 melanoma patients. The selection of serum-reactive clones was based on the biopanning using protein G coated magnetic beads followed by the immunoscreening of the enriched library. Commonly, 4 to 5 rounds of biopanning are performed in order to achieve sufficient enrichment with the phage of interest (Somers et al., 2002; Willats, 2002). We reasoned that in an experiment where polyclonal antibodies with different titres and affinities are used to select potential antigens from a library of phages expressing fusion proteins of different sizes and biochemical properties, performing multiple rounds of biopanning may result in the enrichment of more viable phages recognised by high-titre antibodies and the under-representation of phages whose infectivity or viability is affected by the fusion protein they express. Therefore, initially we assessed the repertoire of antigens selected after the first and the second round of biopanning using two different serum samples. Approximately 5×10^{10} pfu were used for biopanning with MA002079 and LGP-Mel 150 sera that resulted in the recovery of ~5×10^{5} pfu in both cases. Approx. 8×10^{3} pfu of the enriched libraries were subjected to immunoscreening that resulted in the detection of ~80 and 60 serum-reactive clones with MA002079 and LGP-Mel 150 sera,

respectively. The remaining phages were amplified and subjected to the second round of biopanning followed by the immunoscreening with the respective serum. Although a higher enrichment with serum-reactive clones (>130 positive clones per 8×10^3 pfu) was achieved, a considerable reduction in the diversity of the antigen repertoire was observed. The representation of different antigens among clones detected in biopan 1 and 2 fell from 73 to 10% and 92 to 50% when screened with MA002079 and LGP-Mel 150 sera, respectively. Therefore we chose to use a single round of biopanning throughout the study.

Next, serum-reactive clones were selected from T7 phage-displayed testis cDNA library using two more individual sera and a pool of 5 melanoma patients' sera. Approximately 10^4 pfu from the enriched libraries were immunoscreened and that resulted in the detection of 40–150 reactive clones per serum. In total 436 serum-reactive phage clones were isolated, purified via several rounds of immunoscreening and their cDNA inserts were amplified by PCR and sequenced.

3.5. Sequence analysis and characterisation of the identified antigens

The sequence analysis of the identified serum-reactive phage clones revealed that they represent 243 different antigens. However, only in 24 cases cDNAs of known genes were fused in-frame to the phage coat protein 10B thus ensuring that the natural products of these genes are exposed on the surface of the phage (Table 4). Six of these antigens have been previously detected by conventional SEREX, another 10 represent protein families whose other members have been detected by SEREX, and 3 are autoantigens known to

Table 4
In-frame antigens identified by screening T7 phage-displayed testis cDNA library with melanoma patients' sera

Gene symbol	Protein	Serum	Position (aa) on Ref Seq	SEREX DB ID [a]
CCDC84	Coiled-coil domain containing 84	MA002079	186–280, NP_940891.1	–
HDLBP	High density lipoprotein binding protein (vigilin)	MA002079	1161–1268, NP_005327.1	309, gastric Ca
KIF27	Kinesin family member 27	MA002079	715–940 and 715–1128, [b] NP_060046.1	–
MASK-BP3	MASK-4E-BP3 alternate reading frame gene	MA002079	1482–1540, NP_065741.3	1082, renal cell Ca
LIG1	DNA ligase I	Pool-5	1–141, NP_000225	–
LMOD1	Leiomodin 1	Pool-5, MA00GG, LGP-Mel 150 [c]	463–507, NP_036266.2	2374, fibrosarcoma
COPS4	COP9 signalosome subunit 4	Pool-5	15–136, NP_057213.2	–
R3HDM2	R3H domain-containing protein 2	Pool-5	285–513, NP_055740	1233, breast Ca
ANXA11	Annexin A11	Pool-5	1–71, NP_665876.1	81, lung Ca
DKFZP566E164	Hypothetical protein LOC25858	Pool-5	3–76, NP_001034585	–
SENP1	Sentrin/SUMO-specific protease 1	LGP-Mel 143	355–618, NP_055369.1	–
KIF1B	Kinesin family member 1B isoform b	LGP-Mel 143	316–458, NP_055889.2	–
RPS2	Ribosomal protein S2	LGP-Mel 143	263–278, NP_002943.2	182, CTCL, prostate Ca
ATP2C1/ANKRD24	Calcium-transporting ATPase 2C2/ankyrin repeat domain 24	LGP-Mel 143, LGP-Mel 150 [c]	775–785, NP_055676.2/614–650, NP_597732.1	–
SPAG8	Sperm associated antigen 8	LGP-Mel 143, LGP-Mel 150 [c]	9–99, NP_758516.1	–
AKAP12	A-kinase anchor protein 12	LGP-Mel 143, LGP-Mel 150	1465–1622 or 1382–1529, [b] NP_005091.2	–
CCDC92	Coiled-coil domain-containing protein 92 (Limkain beta-2)	LGP-Mel 150	75–200, NP_079416.1	–
ATP5G2	ATP synthase, H+ transporting, mitochondrial F0 complex	LGP-Mel 150	56–107, NP_005167.2	–
LRRC50	Leucine rich repeat containing 50	LGP-Mel 150	626-725, NP_848547.3	–
SLU7	Step II splicing factor	LGP-Mel 150	110–228, NP_006416.3	–
Similar to: LRRC37A	Similar to: leucine rich repeat containing 37A	MA00GG	803–870, NP_055649.3	–
TEF	Thyrotrophic embryonic factor (novel splice variant)	MA00GG	1–52 and 162–215, NP_003207.1	–

[a] Identification number of an antigen in the Cancer Immunome Database (www2.licr.org//CancerImmunomeDB).
[b] Two partially overlapping clones encoding the same antigen were isolated;
[c] The same phage clone was isolated with multiple sera.

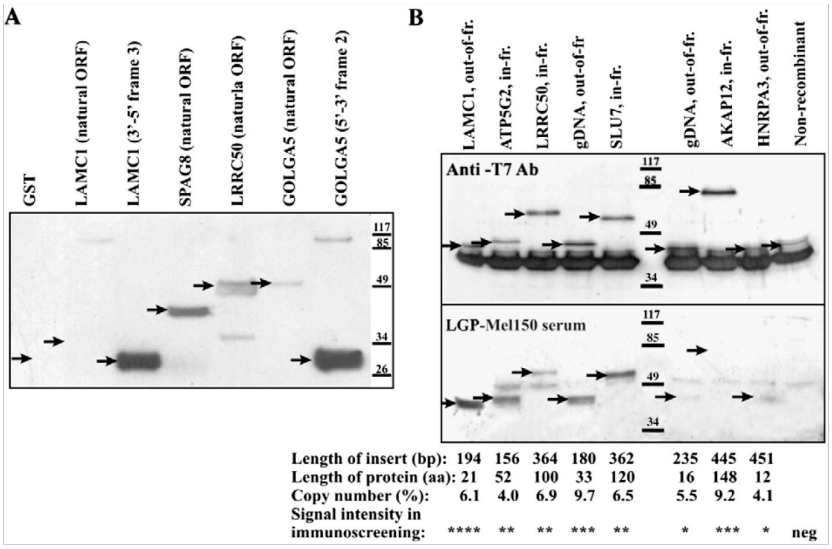


Fig. 3. Western blot analysis of serum reactivity against selected antigens. (A) In-frame and out-of-frame peptides encoded by the selected serum-reactive phages were expressed as GST C-terminal fusion proteins and tested for the reactivity with LGP-Mel 150 serum and anti-GST antibody to confirm the expression of the fusion protein and to determine its size (not shown). The arrows indicate the location of GST fusion proteins. (B) Lysates of ~10^9 pfu of the selected phage particles were separated by SDS/PAGE and tested with anti-10B antibody to determine the copy number of recombinant proteins per phage particle (quantification was done as described in Fig. 1) with patient's serum used for the immunoselection.

induce autoantibody production in autoimmune disorders (LMOD1, AKAP12) or infertility (SPAG8), while no immune responses against 4 antigens — SENP1, TEF, SLU7 and DKFZP566E164 have been reported before. One of the clones contained a hybrid cDNA generated by in-frame fusion of *ATP2C1* and *ANKRD24* genes; however no evidence for the expression of this fusion mRNAs in the testis tissue was found, suggesting that most likely the clone arose as a cloning artefact. Of the remaining 219 clones, 12 represent novel splice variants of known genes or uncharacterised genes for which the natural ORFs have not been determined, 143 represent 5′ or 3′ UTRs or cDNAs fused to 10B in a different reading frame, 52 clones represent intergenic DNA with no corresponding ESTs, 3 clones encode for ribosomal RNA genes and 9 clones contain mitochondrial DNA.

The antibody reactivity against two in-frame antigens (SPAG8 and LRRC50) and two out-of-frame peptides (encoded by *GOLGA5* and *LAMC1*) was confirmed by Western blot analysis. The proteins encoded by the corresponding serum-reactive clones were expressed as GST fusion proteins (for *GOLGA5* and *LAMC1* both — the natural products of these genes and 5′-3′ frame 2 peptide of *GOLGA5* and 3′-5′ frame 3 peptide of *LAMC1* were produced), purified and used for Western blot analysis with anti-GST antibody (not shown) and LGP-Mel 150 serum (Fig. 3A). This confirmed the reactivity against *SPAG8* and *LRRC50* encoded antigens and *GOLGA5* and *LAMC1* out-of-frame peptides.

We further hypothesised that the relatively high proportion of out-of-frame peptides among the serum-reactive clones identified by screening T7 phage-displayed library could be associated with the higher display density of these peptides than natural ORFs due to their shorter size (the size of the out-of-frame peptides ranged from 2 to 56 aa, average 21 aa, while in-frame proteins ranged from 10 to 414 aa, average 120 aa). To test this, a panel of 8 serum-reactive clones (4 in-frame, 4 out-of-frame antigens) and an empty T7 phage was assembled and subjected to Western blot analysis with anti-T7 10B antibody and LGP-Mel 150 serum (Fig. 3B) as described above. The copy number of fusion proteins per phage particle ranged from 17 to 42 (4–10%). No correlation between the copy number and (1) the size of the fusion protein, (2) the size of the insert and (3) signal intensity in Western blot and plaque immunoscreening with patient's serum was observed. This shows that the detection of serum reactivity against out-of-frame peptides is an intrinsic feature of T7 Select 10-3b phage display system that is not associated with the variations in the copy number of recombinant proteins per phage. The serum reactivity was detected against all fusion proteins, except for the AKAP12, hence suggesting that the anti-AKAP12

antibodies recognise a conformational epitope that is absent on denatured protein, while all the other antigens carry linear epitopes.

4. Discussion

In the current study we attempted to evaluate the suitability of T7 Select 10-3b and λKM8 phage display systems for monitoring autoantibody responses against a range of clinically relevant tumour antigens. This showed that both systems are capable of displaying members of all antigen families that were cloned and a relatively good concordance in the detection of autoantibodies by plaque immunoscreening was observed — reactivity against 5 antigen families was detected using both display systems while no reactivity against the remaining 9 antigen families was detected. As the display of the non-reactive antigens at least on the T7 phages was confirmed by Western blot analysis, these are likely to represent true negative calls. However, the sensitivity of the detection of autoantibodies against various antigens seems to differ between the display systems — T7 Select system exhibited a higher sensitivity (in terms of signal intensities and the number of reactive sera) for the detection of autoantibodies against the members of CTAG, MAGEA and GAGE antigen families. In line with the immunoscreening results, T7 phage display system turned out to be more sensitive for the detection of antibodies against CTAG1B/CTAG2 and MAGEA antigens in microarray screening also. This was an unexpected finding, since lambda phage has been shown to be able to assemble capsids consisting of up to ~50% N-terminal fusion protein D (Vaccaro et al., 2006), while T7 phage capsids can incorporate only 1.2–3.6% recombinant proteins according to the manufacturer's description (Novagen) or 1–17.0% according to our results. This suggests that the N-terminus of the recombinant protein D on lambda phage may be spatially less accessible for antibodies than C-terminus of 10B on T7 phage. In fact, the analysis of the crystal structure of gpD demonstrated that the N-termini up to Ser 15 are disordered and are located near the three-fold axis of gpD trimer on the side that binds to the capsid surface and hence at least partially may be hidden under the gpD trimer (Yang et al., 2000).

At the same time, the reactivity against CTAGE5 could be detected only when it was displayed on lambda phage but not on T7 phage. Since the copy number of fusion proteins on CTAGE5 encoding T7 phages was the lowest among all the phages analysed (~1%), most likely, it did not reach the detection limit thus preventing the serum reactivity it to be detected. Hence, the very variable display density of T7 Select system confers a risk of false negative calls due to an insufficient copy number and furthermore suggests that the signal intensity in plaque assay or microarrays may depend not only on the antibody titre but also on the copy number of recombinant proteins. However, this disadvantage could be overcome by constructing a novel vector that would allow monitoring the copy number of fusion proteins per phage particle.

Next, we applied T7 Select 10-3b phage display-based SEREX approach to search for antigens eliciting immune responses in melanoma patients that resulted in the identification of 436 serum-reactive clones representing 243 different antigens. However, only 24 of them represented known genes translated in their natural reading frames and another 12 were novel splice variants or uncharacterised genes (with at least two ESTs confirming that these sequences are transcribed) with unknown protein sequences. Six of the in-frame antigens have been previously detected by conventional SEREX and another 10 represent protein families whose other members have been detected by SEREX but no immune responses to 4 antigens have been reported before thus demonstrating that the repertoire of antigens identified by T7 phage display-based SEREX approach overlaps with conventional SEREX, at the same time may allow the detection of novel antigens.

Some of them have been previously shown to have a diagnostic value or may play a significant role in cancer progression. For example, the presence of autoantibodies against Annexin XI-A could significantly discriminate between breast cancer and non-cancer control sera and their frequency was higher in patients with ductal carcinoma *in situ* than in invasive ductal carcinoma (Fernandez-Madrid et al., 1999). Another of the identified antigens, sperm-associated antigen 8 (SPAG8), so far was implicated in a rare form of female infertility, where anti-SPAG8 antibodies have been shown to cause sperm agglutination (Zhang et al., 2000). This protein was shown to be predominantly expressed on the acrosome of sperm and functionally involved in the acrosome reaction and sperm binding to the zona pellucida (Cheng et al., 2007). Very recently it was found to be overexpressed in HPV18 infected cervical cancer cells (Vazquez-Ortiz et al., 2007). At the same time, a closely related protein, SPAG1, was shown to be expressed at high levels in a large proportion of pancreatic ductal adenocarcinomas and contribute to cancer cell motility. Hence, SPAG may represent a novel CT antigen family, however more detailed analysis of their expression in normal and cancerous tissues is required. Moreover, considering its localisation on cell surface at least on spermatozoa, it seems to be an attractive target for antibody-based therapeutical approaches.

The remaining 219 clones contained fragments of intergenic DNA (52), mtDNA (9), rRNA (3), 5′or 3′ UTRs or cDNAs (143) cloned out-of-frame relatively to the coat

protein 10B. It cannot be excluded that some of the 143 clones encoding UTRs or cDNAs in alternative ORFs indeed represent cancer-specific antigens generated by frame-shifting mutations, defects in pre-mRNA splicing or aberrations in translational controls in cancer cells, as evidenced by the detection of serum-reactive clones expressing CTAG1B-ORF2 peptide by screening TAA mini-libraries. Nevertheless, most likely the majority of these clones, particularly those 52 clones containing intergenic regions with no evidence of expression, display peptides that are not naturally expressed and therefore can be considered as mimotopes. A similar proportion of in-frame and out-of-frame antigens has been found by Chatterjee et al. (2006) in a study where serum-reactive clones were selected from T7 phage-displayed ovarian cancer cDNA library. Out-of-frame peptides, yet in smaller number, also have been detected by using pJuFo phage display system (Fossa et al., 2004) and lambda phage surface display (Minenkova et al., 2003). The nature of the antigens they mimic is not known and we believe — cannot be unambiguously determined by BLAST search through protein databases because the exact epitope sequences are not known and, moreover, they may mimic protein as well as non-protein antigens of various pathogens not only cancer cells. Consequently, the probability of finding cancer-associated biomarkers among them should be lower than among the in-frame antigens.

Experience of our and other groups (personal communication Dr. G. Li; Nottingham and Dr. S. Eichmuller, Heidelberg) has shown that approximately 1/3 of the serum-reactive clones identified by conventional SEREX also contain cDNAs fused to β-galactosidase gene in a non-natural reading frame. Although it is possible that the natural products of these genes could be produced by means of using alternative ribosomal binding sites, to our knowledge it has never been experimentally confirmed. Therefore we assume that a considerable fraction of these antigens also represents mimotopes, though their incidence is much lower than in T7 phage display-based SEREX approach. We reasoned, the identification of so high percentage of mimotopes using T7 phage could be caused by a higher display density of the out-of-frame peptides and subsequent detection of low affinity antibodies due to a higher valence of the epitope carrier. However, the analysis of the display density on the selected phage particles did not confirm this hypothesis, hence either the overrepresentation of the phages expressing shorter peptides in the amplified libraries or the folding properties and accessibility of the C-terminus of the coat protein 10B are likely to be responsible for this phenomenon.

The application of phage display technology to cancer serology offers time-, labour- and cost-effective alternative to conventional SEREX. Furthermore, it gives an opportunity to produce antigen chips by printing the recombinant phage particles on microarray slides thus allowing avoidance of production and purification of recombinant proteins that would not be feasible for all the antigens identified. This in turn enables the analysis of the whole autoantibody profile in patients' sera that would allow establishing the significance of the autoantibody profiles, not individual autoantibodies, as biomarkers for the early detection and prognosis of cancer and prediction of response to immunotherapy. So far, T7 phage, presumably due to its favourable biological properties (fast growing, chemically resistant and easy to obtain high-titre stocks) and availability of good commercial antibodies against the phage tail protein, has been the vector of choice for the production of antigen microarrays (Fernandez-Madrid et al., 1999; Zhong et al., 2005; Wang et al., 2005; Chatterjee et al., 2006). However, the variable copy number and the display of very high percentage of mimotopes constitute the two main drawbacks for exploiting T7 Select system for the analysis of autoantibody profiles. Here we demonstrated that the lambda phage is equally suitable for the production of phage-displayed antigen microarrays — it turned out to be possible to obtain high-titre phage stocks without any concentration or purification steps, the phage capsid appeared to be sufficiently stable and in the most cases the displayed proteins retained their capacity to be recognised by antibodies, however it was less sensitive than T7 phage for the detection of antibodies against several antigens. Nevertheless, as the lambda phage is capable to assemble capsids with N-terminal gpD fusion proteins that efficiently diminish display of out-of-frame peptides, it could be a preferential display system for the studies aiming to define novel potential therapeutic targets or to assess the presence of autoantibodies against known tumour antigens.

In conclusion, the exploitation of phage display-based approaches for the identification of tumour antigens provides a time- and labour-effective alternative to the conventional SEREX allowing the identification of a diverse antigen repertoire that partially overlaps with SEREX. Moreover, both T7 Select and λKM8 display systems are equally suitable for the production of antigen microarrays allowing the monitoring of autoantibody profiles, however they differ in the sensitivity of the detection of antibodies against various antigens.

Acknowledgments

We are thankful to Dr. Mauruzio Cianfriglia for providing anti-gpV monoclonal antibody and to Dr.

Aivars Stengrēvics for providing a part of the collection of sera from melanoma patients. We appreciate the help of Antje Sucker and Dr. Stefan Eichmüller for selecting appropriate serum samples.

This study was supported by EU 6th Framework Program ENACT (LSHC-CT-2004-503306) and a grant from Latvian State Research Program No. 07-VP-2, and fellowships from ESF.

References

Bachelot, T., Ratel, D., Menetrier-Caux, C., Wion, D., Blay, J.Y., Berger, F., 2006. Autoantibodies to endostatin in patients with breast cancer: correlation to endostatin levels and clinical outcome. Br. J. Cancer 94, 1066.

Cekaite, L., Haug, O., Myklebost, O., Aldrin, M., Ostenstad, B., Holden, M., Frigessi, A., Hovig, E., Sioud, M., 2004. Analysis of the humoral immune response to immunoselected phage-displayed peptides by a microarray-based method. Proteomics. 4, 2572.

Chatterjee, M., Mohapatra, S., Ionan, A., Bawa, G., li-Fehmi, R., Wang, X., Nowak, J., Ye, B., Nahhas, F.A., Lu, K., Witkin, S.S., Fishman, D., Munkarah, A., Morris, R., Levin, N.K., Shirley, N.N., Tromp, G., Abrams, J., Draghici, S., Tainsky, M.A., 2006. Diagnostic markers of ovarian cancer by high-throughput antigen cloning and detection on arrays. Cancer Res. 66, 1181.

Cheng, G.Y., Shi, J.L., Wang, M., Hu, Y.Q., Liu, C.M., Wang, Y.F., Xu, C., 2007. Inhibition of mouse acrosome reaction and sperm-zona pellucida binding by anti-human sperm membrane protein 1 antibody. Asian J. Androl. 9, 23.

Fernandez-Madrid, F., VandeVord, P.J., Yang, X., Karvonen, R.L., Simpson, P.M., Kraut, M.J., Granda, J.L., Tomkiel, J.E., 1999. Antinuclear antibodies as potential markers of lung cancer. Clin. Cancer Res. 5, 1393.

Fossa, A., Alsoe, L., Crameri, R., Funderud, S., Gaudernack, G., Smeland, E.B., 2004. Serological cloning of cancer/testis antigens expressed in prostate cancer using cDNA phage surface display. Cancer Immunol. Immunother. 53, 431.

Hansen, M.H., Ostenstad, B., Sioud, M., 2001. Identification of immunogenic antigens using a phage-displayed cDNA library from an invasive ductal breast carcinoma tumour. Int. J. Oncol. 19, 1303.

Hufton, S.E., Moerkerk, P.T., Meulemans, E.V., de Bruine, A., Arends, J.W., Hoogenboom, H.R., 1999. Phage display of cDNA repertoires: the pVI display system and its applications for the selection of immunogenic ligands. J. Immunol. Methods 231, 39.

Krumpe, L.R., Atkinson, A.J., Smythers, G.W., Kandel, A., Schumacher, K.M., McMahon, J.B., Makowski, L., Mori, T., 2006. T7 lytic phage-displayed peptide libraries exhibit less sequence bias than M13 filamentous phage-displayed peptide libraries. Proteomics 6, 4210.

Li, Y., Karjalainen, A., Koskinen, H., Hemminki, K., Vainio, H., Shnaidman, M., Ying, Z., Pukkala, E., Brandt-Rauf, P.W., 2005. p53 autoantibodies predict subsequent development of cancer. Int. J. Cancer 114, 157.

Lucchese, A., Willers, J., Mittelman, A., Kanduc, D., Dummer, R., 2005. Proteomic scan for tyrosinase peptide antigenic pattern in vitiligo and melanoma: role of sequence similarity and HLA-DR1 affinity. J. Immunol. 175, 7009.

Minenkova, O., Pucci, A., Pavoni, E., De, T.A., Fortugno, P., Gargano, N., Cianfriglia, M., Barca, S., De, P.S., Martignetti, A., Felici, F., Cortese, R., Monaci, P., 2003. Identification of tumor-associated antigens by screening phage-displayed human cDNA libraries with sera from tumor patients. Int. J. Cancer 106, 534.

Pallasch, C.P., Struss, A.K., Munnia, A., Konig, J., Steudel, W.I., Fischer, U., Meese, E., 2005. Autoantibodies against GLEA2 and PHF3 in glioblastoma: tumor-associated autoantibodies correlated with prolonged survival. Int. J. Cancer 117, 456.

Pavoni, E., Vaccaro, P., Pucci, A., Monteriu, G., Beghetto, E., Barca, S., Dupuis, M.L., De Pasquale, C.A., Lugini, A., Cianfriglia, M., Cortesi, E., Felici, F., Minenkova, O., 2004. Identification of a panel of tumor-associated antigens from breast carcinoma cell lines, solid tumors and testis cDNA libraries displayed on lambda phage. BMC Cancer 4, 78.

Rosenberg, A., Griffin, K., Studier, W.F., McCormick, M., Berg, J., Novy, R., Mierendorf, R., 1996. T7 Select Phage Display System: a powerful new protein display system based on bacteriophage T7. Innovations 1–6.

Sahin, U., Tureci, O., Schmitt, H., Cochlovius, B., Johannes, T., Schmits, R., Stenner, F., Luo, G., Schobert, I., Pfreundschuh, M., 1995. Human neoplasms elicit multiple specific immune responses in the autologous host. Proc. Natl. Acad. Sci. U. S. A 92, 11810.

Sioud, M., Hansen, M.H., 2001. Profiling the immune response in patients with breast cancer by phage-displayed cDNA libraries. Eur. J. Immunol. 31, 716.

Slager, E.H., Borghi, M., van der Minne, C.E., Aarnoudse, C.A., Havenga, M.J., Schrier, P.I., Osanto, S., Griffioen, M., 2003. CD4+ Th2 cell recognition of HLA-DR-restricted epitopes derived from CAMEL: a tumor antigen translated in an alternative open reading frame. J. Immunol. 170, 1490.

Somers, V.A., Brandwijk, R.J., Joosten, B., Moerkerk, P.T., Arends, J.W., Menheere, P., Pieterse, W.O., Claessen, A., Scheper, R.J., Hoogenboom, H.R., Hufton, S.E., 2002. A panel of candidate tumor antigens in colorectal cancer revealed by the serological selection of a phage displayed cDNA expression library. J. Immunol. 169, 2772.

Soussi, T., 2000. p53 Antibodies in the sera of patients with various types of cancer: a review. Cancer Res. 60, 1777.

Stockert, E., Jager, E., Chen, Y.T., Scanlan, M.J., Gout, I., Karbach, J., Arand, M., Knuth, A., Old, L.J., 1998. A survey of the humoral immune response of cancer patients to a panel of human tumor antigens. J. Exp. Med. 187, 1349.

Usener, D., Schadendorf, D., Koch, J., Dubel, S., Eichmuller, S., 2003. cTAGE: a cutaneous T cell lymphoma associated antigen family with tumor-specific splicing. J. Invest. Dermatol. 121, 198.

Vaccaro, P., Pavoni, E., Monteriu, G., Andrea, P., Felici, F., Minenkova, O., 2006. Efficient display of scFv antibodies on bacteriophage lambda. J. Immunol. Methods 310, 149.

Vazquez-Ortiz, G., Garcia, J.A., Ciudad, C.J., Noe, V., Penuelas, S., Lopez-Romero, R., Mendoza-Lorenzo, P., Pina-Sanchez, P., Salcedo, M., 2007. Differentially expressed genes between high-risk human papillomavirus types in human cervical cancer cells. Int. J. Gynecol. Cancer 17, 484.

Wang, X., Yu, J., Sreekumar, A., Varambally, S., Shen, R., Giacherio, D., Mehra, R., Montie, J.E., Pienta, K.J., Sanda, M.G., Kantoff, P.W., Rubin, M.A., Wei, J.T., Ghosh, D., Chinnaiyan, A.M., 2005. Autoantibody signatures in prostate cancer. N. Engl. J. Med. 353, 1224.

Welling, G.W., Weijer, W.J., van der Zee, R., Welling-Wester, S., 1985. Prediction of sequential antigenic regions in proteins. FEBS Lett. 188, 215.

Willats, W.G., 2002. Phage display: practicalities and prospects. Plant Mol. Biol. 50, 837.

Yang, F., Forrer, P., Dauter, Z., Conway, J.F., Cheng, N., Cerritelli, M.E., Steven, A.C., Pluckthun, A., Wlodawer, A., 2000. Novel fold and

capsid-binding properties of the lambda-phage display platform protein gpD. Nat. Struct. Biol. 7, 230.

Zhang, X.D., Miao, S.Y., Wang, L.F., Li, Y., Zong, S.D., Yan, Y.C., Koide, S.S., 2000. Human sperm membrane protein (hSMP-1): a developmental testis-specific component during germ cell differentiation. Arch. Androl. 45, 239.

Zhong, L., Hidalgo, G.E., Stromberg, A.J., Khattar, N.H., Jett, J.R., Hirschowitz, E.A., 2005. Using protein microarray as a diagnostic assay for non-small cell lung cancer. Am. J. Respir. Crit Care Med. 172, 1308.

Zhong, L., Peng, X., Hidalgo, G.E., Doherty, D.E., Stromberg, A.J., Hirschowitz, E.A., 2004. Identification of circulating antibodies to tumor-associated proteins for combined use as markers of non-small cell lung cancer. Proteomics 4, 1216.

Zippelius, A., Gati, A., Bartnick, T., Walton, S., Odermatt, B., Jaeger, E., Dummer, R., Urosevic, M., Filonenko, V., Osanai, K., Moch, H., Chen, Y.T., Old, L.J., Knuth, A., Jaeger, D., 2007. Melanocyte differentiation antigen RAB38/NY-MEL-1 induces frequent antibody responses exclusively in melanoma patients. Cancer Immunol. Immunother. 56, 249.

Zucconi, A., Dente, L., Santonico, E., Castagnoli, L., Cesareni, G., 2001. Selection of ligands by panning of domain libraries displayed on phage lambda reveals new potential partners of synaptojanin 1. J. Mol. Biol. 307, 1329.

3.2 Analyses of the identified natural and undefined ORF cancer antigens

Molecular analysis of serologically defined tumor antigens reveals novel cancer-testis antigens as potential targets for cancer immunotherapy.

Karīna Siliņa[1], Pavel Zayakin[1], Zane Kalniņa[1], Lāsma Ivanova[1], Irēna Meistere[1], Aivars Stengrēvics[2], Thomas Wex[4], Else Marit Inderberg Suso[3], Mārcis Leja[2], Peter Malfertheiner[4], Gustav Gaudernack[3], Dirk Schadendorf[5], Aija Linē[1]

1 Latvian Biomedical Research and Study Centre, Ratsupites str. 1, Riga, LV-1067, Latvia
2 Riga Eastern Clinical University Hospital, Hipokrata str. 2, Riga, Latvia
3 Department of Immunology, Institute for Cancer Research Oslo University Hospital, Montebello, 0310 Oslo, Norway
4 Clinic of Gastroenterology, Hepatology and Infectious Diseases, Otto-von-Guericke University Magdeburg, Leipziger Str. 44, Magdeburg, 39120, Germany
5 Department of Dermatology, University Hospital Essen, Hufelandstraße 55, 45147 Essen, Germany

Correspondence to:
Karina Silina
karina@biomed.lu.lv
Latvian Biomedical Research and Study Centre
Ratsupites 1
Riga
LV-1067
Latvia

ABSTRACT

Previously using phage display-based SEREX approach we identified 1328 antigens recognised by autoantibodies from melanoma, breast, prostate and gastric cancer patients. 198 antigens were translated in their natural open reading frames (ORF) including well-known cancer-testis (CT) and known autoantigens as well as 79 previously uncharacterized antigens, 21 antigens were encoded by uncharacterized genes or novel splice variants termed as undefined ORF antigens, while the remaining clones were translated as unnatural ORF peptides. The current study was aimed to find out whether any of the novel natural and undefined ORF antigens could serve as targets for cancer immunotherapy. First, their autoantibody frequency was assessed in sera from 376 cancer patients and 153 healthy individuals using phage-displayed antigen microarrays. Next, antigens with cancer-related humoral response were prioritized according their EST tissue distribution, functional significance and structural motifs and 30 top-ranked antigens were subjected to mRNA expression analysis in 14 various normal tissues and melanoma, breast, and gastric cancer-normal tissue pairs. Normal tissue analysis revealed eight testis-selective and four testis-restricted transcripts. Two of them represented novel splice variants of CFL1 and COX6B2 and were detectable in cancer samples suggesting that deregulated alternative splicing in cancer can generate germ cell-associated isoforms contributing to cancer antigenicity. Another four testis-selective antigens – ALLC, C11ORF20, C16ORF82, and LOC284861, were upregulated in ~10% of cancer specimens, while two antigens – LRRC50 and ESCO1, were significantly overexpressed in 25-60% of various cancers revealing them as novel members of the CT antigen category with potential application in immunotherapy.

Keywords: CT antigen, CT-spliced antigen, phage-display SEREX, antigen microarray, cancer immunotherapy

INTRODUCTION

The genetic and epigenetic instability of cancer cells underlies not only their malignant phenotype, but also their immunogenicity creating tumor-specific neo-antigens characterized by structural alterations of proteins and tumor-associated antigens characterized by ectopic or elevated expression in tumours. The structural changes are introduced through mutations, chromosomal translocations, altered protein modifications as well as splicing defects and aberrant translation regulation. The latter two mechanisms can create the so-called unconventional antigens and include cryptic peptides that are formed by exon/intron junctions, translation of alternative open reading frames (ORF), subdominant ORFs located in the 3'or 5' untranslated regions (UTR) of cDNAs, and intronic sequences [1]. Tumor-associated antigens are created by aberrant expression of genes normally expressed only in immunoprivileged sites and fetal tissues such as the cancer-testis (CT) and oncofetal antigens, terminally differentiated particular cell types as the differentiation antigens, or by overexpression of genes that are normally expressed in much less amount than in tumor cells as the overexpressed antigens [2].

Tumor immunotherapy offers an appealing strategy to conquer cancers bearing such antigens in a specific and non-toxic manner, and promising therapy targets have been found in both, qualitatively and quantitatively altered antigen groups [3]. However, the documented overall success rate is low due to factors that subvert the therapy-induced immune response like the activation of innate or adaptive tolerogenic immune cells and the selection of tumor escape variants through the downregulation of the elements of antigen presentation pathway, the loss of antigen expression and the production of immunosuppressive cytokines by tumor or its stroma [4, 5]. The search for protective antigens is still a burning task in order (i) to develop polyepitopic targeting to diminish the possibility of tumor escape variant selection through antigen downregulation, and to cover for the antigenic heterogeneity of each tumor as well as for the tumors that don't express the currently acknowledged target antigens, and (ii) to develop tools for monitoring immunotherapy induced anti-tumor T cell response.

Recently guidelines of prioritizing antigens as targets for immunotherapy have been elaborated by a wide panel of experts [3] defining that an ideal cancer antigen is such that is effective in immunotherapy, immunogenic, specific to cancer cells, oncogenic, highly present in all cancer cells including putative cancer stem cells, present in tumors of many patients, multiepitopic, and is located on cell surface. The SEREX technology allows to identify proteins that have elicited the formation of class switched antibodies, which may indicate to spontaneous T helper cell activation against the particular antigens[6]. Tumor antigen-targeted immunotherapy trials have shown that the expansion of antigen-specific cytotoxic T lymphocytes does not necessarily correlate with tumor regression [7], while the proper activation of antigen specific T helper cells has a great potential to eradicate tumors and might be the missing piece in the successful immunotherapy puzzle [8]. We applied phage-displayed SEREX in our previous study to identify a comprehensive set of antigens eliciting humoral responses in melanoma, gastric, breast and prostate cancer that could be used as cancer serum biomarkers for diagnosis, prognosis and prediction of immunotherapy outcome, and this resulted in the collection of over 1300 antigenic clones ([9] and unpublished results). The objective of the present study was to determine whether any of the identified antigens could serve as novel immunotherapeutic targets matching the above criteria.

MATERIALS AND METHODS

Patient material and RNAs

Tumor and adjacent histologically normal tissue specimens were obtained from operation material of melanoma, gastric and breast cancer patients undergoing surgery in Latvian Oncology Centre and stored in RNALater® (Applied Biosystems, USA). Tissue sections were evaluated by pathologists and clinical information was collected. The specimens were collected after the patients' informed consent was obtained in accordance with the regulations of Central Commission of Medical Ethics of Latvia.

Normal human tissue RNA panels were purchased from Applied Biosystems, USA and Biocat, Germany. RNA from melanoma cell lines was kindly provided by Skin Cancer Unit in German Cancer Research Center.

Sera were collected from the same patients whose tissue specimens were collected. Additional serum samples of breast cancer patients and healthy individuals with no known history of cancer and autoimmune disorders were provided by the Genome Database of the Latvian population and sera from melanoma and gastric cancer patients were kindly provided by the Skin Cancer Unit in German Cancer Research Center, the Clinic of Gastroenterology, Hepatology and Infectious Diseases, Otto-von-Guericke University Magdeburg, and Norwegian Radium Hospital.

Autoantibody profiling assay

Production of phage-displayed antigen microarray, its processing and data analysis is described elsewhere (Zayakin P et al., article in preparation, 2011). Briefly, a custom antigen microarray comprising 1158 recombinant and 71 non-recombinant T7 phage clones was produced as described previously [9] and was screened with sera from 190 melanoma, 173 gastric cancer, 13 breast cancer, and 153 healthy individuals. The microarray slides were incubated with patients' sera at dilution of 1:200 and with a mouse anti-T7 phage tail antibody (Novagen, USA). The serum reactivity was detected by Cy5 conjugated goat anti-human IgG antibody (Jackson ImmunoResearch, USA) and the ratio against the total amount of printed phage, which was detected by Cy3 conjugated goat anti-mouse IgG antibody (Jackson ImmunoResearch, USA) was calculated. A two-step normalization strategy was used for the fluorescent signal ratios in order to eliminate variations introduced by custom production of microarrays and variable background reactivity of different sera. At first, the values in each slide (each serum) were normalized by the median of all printed spots for each fluorescent channel separately. Next, the Cy5 and Cy3 signal intensities for each spot were divided by the median value of that spot within the print lot. For autoantibody frequency calculations, the cutoff value of specific seroreactivity was defined as four standard deviations above the average of 71 non-recombinant phage controls. Statistical significance was calculated using χ^2 test.

In Silico analysis

To determine candidate CT or overexpressed antigens as well as structurally altered unconventional antigens, the following bioinformatics tools were exploited. For the analysis of EST tissue distribution Human EST database was downloaded from ftp://ftp.ncbi.nih.gov/genomes/ H_sapiens/ and each clone sequence was aligned against EST database using megablast algorithm, EST sequences with E-value <1e-30, >70% coverage and >97% identities were included in the further analysis. Repeat sequences were masked by applying filters for low complexity regions and human specific repeats. ESTs were classified as "cancer", "cell line", "normal", "embryonic", "germ cell" and "brain" by extracting tissue source information from their annotations. The EST frequency for each

antigen clone per category was calculated relatively to the total number of ESTs in that category. Those clones whose EST frequency in cancer and/or cell line categories was two-fold higher than in normal tissue category were considered to be predominantly expressed in cancers, while those that showed 2-fold higher frequency in germ cells and/or embryonic tissue categories than in normal tissues were considered to be predominantly expressed in germ cells, and those with a 10-fold higher frequency in cancer and germ cell categories than in normal tissues were categorized as overexpressed in cancer and germ cells.

For the prediction of protein structure motifs and coiled-coil domains MOTIF Search (http://motif.genome.jp/) and http://www.russell.embl-heidelberg.de tools were used. Analysis of alternative splicing isoforms and splice site prediction was performed using Spiday tool (http://www.ncbi.nlm.nih.gov/spidey/) and putative alternative ORFs were predicted using Translate tool at ExPASy Proteomics server (http://www.expasy.org/tools/dna.html). NCBI protein BLAST was used to search for possible sequence homologies of undefined ORF antigens representing novel genes.

RNA extraction, cDNA synthesis

Bead-based tissue homogenization was performed by using the FastPrep-24 instrument and Lysing Matrix A (MP Biomedicals, USA) in 1ml of TRIsolution® (Applied Biosystems, USA) followed by the extraction of total RNA according to manufacturer's protocol. RNA extracted from tissue material was treated with DNAse prior to cDNA synthesis according to manufacturer's protocol (Applied Biosystems, USA). cDNA was synthesized by random hexamer priming from 2µg of total RNA by using RevertAidTM First StrandcDNA Synthesis Kit (Fermentas, Lithuania) according to manufacturer's instructions.

mRNA expression analysis

Qualitative RT-PCR reaction mixtures contained 1x reaction buffer, 2.5mM $MgCl_2$, 0.1µM primers, 0.75U Taq DNA polymerase (Fermentas, Lithuania), and $1/60^{th}$ of the prepared cDNA. Primers were designed to amplify the identified antigenic regions of analysed genes (Table S1). Amplification of all target and reference genes was performed at the same cycling conditions (30s at 94°C, 30s at 60°C, 30s at 72°C), except for the number of cycles that was adjusted individually according to corresponding mRNA abundance.

Quantitative RT-PCR (qPCR) reactions were performed using $1/60^{th}$ of cDNA reaction mixtures, ABSolute Blue™ SYBR green Low ROX (Thermo Scientific, USA) on ABI7500 sequence detection system (Applied Biosystems). Appropriate primer concentrations were established by cDNA 4 log serial dilution curves to ensure amplification efficiency over 95%. Three most stable reference genes were determined among 7 most often used house keeping genes (GAPDH, ACTB, POLR2A, TUB3A, TBP, YWHAZ, PGK1) by using geNorm software for the set of various normal tissues, as well as for the sets of tumor-normal tissue pairs of each cancer type. To normalize the expression data a normalization factor was calculated for each cDNA from the expression values of the established stable reference genes that were (i) ACTB, POLR2A, TUB3A in the various normal tissue set, (ii) YWHAZ, ACTB, TBP in the gastric cancer set, (iii) YWHAZ, TBP, POLR2A in the melanoma set, and (iv) YWHAZ, PGK1, GAPDH in the breast cancer set. All reactions were performed in duplicates. The expression level of each gene was determined relative to its expression in testis. Statistical analysis of expression data was performed by the non-parametric Mann-Whitney U test and Wilcoxon signed rank test.

RESULTS

Description of antigen clone collection

In our previous study aimed to define a repertoire of cancer antigens eliciting humoral responses in patients, T7 phage displayed cDNA expression libraries were constructed from 5 melanoma specimens, 11 gastric cancer specimens, 3 prostate cancer cell lines and 2 testis specimens. Phage clones recognized by serum IgGs were selected from these libraries via biopanning followed by immunoscreening using sera from 26 melanoma, 22 gastric cancer, 22 prostate cancer, one breast cancer and 15 gastritis patients that resulted in the identification of 1328 different phage clones representing 1158 non-redundant antigens ([9] and unpublished data). The sequence analysis of serum reactive clones showed that 60% of the collection is represented by cDNA clones, while 40% appear to be non-cDNA inserts including rRNA and, despite the vigorous DNAse treatment of mRNA before cloning, also intergenic regions, mtDNA, and microbial DNA fragments. In the applied vector system the insert is cloned as the 3' fusion to the T7 phage surface protein 10B, hence both, fragments translated in their natural as well as unnatural ORFs can be displayed. The sequence analysis of recombinant peptides showed that 15% (198) of the clones were translated in their natural ORF, 2% (21) antigens were encoded by uncharacterized genes or novel transcript variants termed here as undefined ORF antigens, and 83% (1109) were clones translated as unnatural ORF peptides absent from any human protein.

Natural ORF antigens represented 130 non-redundant antigens comprising 46 well-characterized CT antigens like the members of CTAG, MAGE etc. families, 5 autoantigens such as ANXA11 [10], and AKAP12 [11] known to induce antibody production in autoimmune disorders, 20 antigens that have been previously identified by conventional SEREX, 34 antigens were from protein families whose other members have been detected by SEREX, and 25 were novel antigens against which no immune response has been reported before. We didn't detect mutations in any of these cDNAs, but we found novel alternative splice variants (ASVs) for three genes, KIF27 (HM370398), TEF (HM370396), and UBR2 (HM370397), that were not annotated in the NCBI AceView database [12].

We introduced a category of antigens called the undefined ORF antigens that encompassed 13 clones representing uncharacterised ASVs translated as alternative or subdominant ORFs whose generation was predicted as theoretically possible by our bioinformatics approach that included search of translation initiation codons of the corresponding ORFs in alternative upstream exon combinations and predictions of internal ribosome entry sites (IRES) (see Materials and Methods). Also 8 clones representing putative novel genes supported by ESTs or a conventional splicing event and for which the ORF has not been characterised were attributed to the undefined ORF category.

Selection of candidate antigens

In order to identify novel potential therapy targets we further analysed only uncharacterised natural ORF antigens for which no comprehensive immunogenicity or expression data existed as well as undefined ORF antigens. To select the most promising candidates for further experimental analysis we ranked the novel natural ORF antigens using the following prioritization criteria: (i) frequency of cancer-related autoantibody response, (ii) CT-associated EST profile, (iii) functional role in oncogenesis, and (iv) structural features as cell surface localization or novel ASVs.

To identify antigens with cancer-related autoantibody response we screened our custom phage-displayed antigen microarray comprising all non-redundant antigens (1158) and 71 non-recombinant phage clones as negative controls with sera from 173 gastric

cancer, 13 breast cancer, 190 melanoma patients, and 153 healthy individuals. The production of the microarray, data normalization and processing is described elsewhere ([9] and Zayakin P et al., article in preparation, 2011). A cutoff value of 4 SDs above the mean serum signal intensity of non-recombinant phage spots was found to clearly discriminate between the presence and absence of specific autoantibodies to most of the antigens. This cutoff was validated experimentally by testing the reactivity of two CTAG1B reactive sera (signal intensities over 5 and 7 SDs above the mean signal intensity of non-recombinant phages) and two non-reactive sera against the CTAG1B by Western blot analysis. Antigens were classified as cancer-related and cancer-nonrelated according to their seroreactivity in each tumor type using cancer versus normal odds ratio (OR)>3.5 as the discriminator (Table 1). The cancer-related antigens were further divided into four subcategories by the autoantibody frequency and ranks from 0 to 3 were assigned for each category as follows: (i) 3 – frequent, statistically significant antigens, (ii) 2 – moderately frequent, putatively cancer-specific antigens, (iii) 1 – cancer-associated antigens, and (iv) 0 – rare, putatively cancer-specific antigens (Table 1). The latter group contains also the antigens that did not react with any serum in the tested set in order not to leave out the antigens that elicit rare, yet cancer-specific humoral response. Cancer-related autoantibody response was detected for 50 uncharacterised natural ORF antigens and nine undefined ORF antigens comprising five novel genes and four novel ASVs.

To determine the antigens with CT-associated EST profile the tissue distribution of all ESTs complementary to a particular antigen was characterised using the information in the EST sequence annotation, and we established categories of EST distribution as described in Materials and Methods. Ranks were assigned to these categories as follows: (i) 3 – overexpressed in cancer and germ cells, (ii) 2 – predominant expression in germ cells or embryonic tissue, (iii) 1 – predominant expression in cancer or cell lines, (iv) 0 – no association to cancer or germ cells.

To discern antigens with potential implications in oncogenesis we searched the published literature and introduced three categories that were ranked as follows: (i) 2 – known oncogenic function or potentially related to oncogenesis or stem cell biology with published circumstantial evidence, (ii) 1 – unknown relation to oncogenesis, (iii) 0 – known function, no relation to oncogenesis. As well such structural properties as the cell surface localization or novel ASV gave additional one point, while a point was subtracted if the antigen contained a coiled-coil domain as these structures are often recognised by cross-reacting antibodies.

The total score obtained as the sum of the above criteria was used to rank the natural ORF antigens and the top 23 antigens were selected for further mRNA expression analysis (Table 2). The criteria for undefined ORF antigen selection were (i) the cancer-related serum response and (ii) no matching ESTs or testis and/or cancer-associated EST profile, and two putative novel splice variants and five putative novel genes were selected for further experiments (Table 3). The description of clone structure, translated peptides together with homology results from protein BLAST, and chromosomal localization of these antigens are indicated in Table 3.

mRNA tissue distribution of selected antigens

Natural ORF antigens

mRNA expression of the selected 23 natural ORF antigens was analysed in sets of various normal tissues and tumour-normal tissue pairs. Hofmann et al. has introduced terms to describe the expression pattern of CT genes and has separated them into two groups: genes that are expressed mostly in testis but are detected in low levels in other non-germ line tissues are called testis-selective, and the ones expressed only in germ line

tissues and placenta are called testis-restricted [13]. All of the genes were initially analyzed by qualitative RT-PCR in 11 normal tissues and the ones that appeared testis-associated, were further verified by quantitative real time PCR (qPCR) in 14 various normal tissues. As a result six testis-associated genes were identified. ALLC, LRRC50, C11ORF20 and ESCO1 were detected mainly in testis, but also in some other non-germ line tissues, but qPCR data demonstrated that they together with C16ORF82, which initially appeared testis-restricted, can be ascribed to the testis-selective gene group (Fig. 1). C21ORF66 was detected exclusively in the testis sample and is a testis-restricted gene. Quantitative mRNA expression analysis didn't confirm testis-selectivity for COPS4 and CCDC92 as appeared in the qualitative RT-PCR (Fig. 1a, b). One of the antigens – a novel ASV of TEF (here designated as TEF-ASV), is translated in its natural ORF, but lacks 105 amino acids from the natural protein caused by an alternative exon skipping event. This transcript was cloned from a testis cDNA library, however we failed to detect its expression in any of the analysed tissues including testis samples from four different individuals. The remaining 14 genes were ubiquitously expressed (Fig. 1a).

Next, all the antigens were quantitatively analyzed in sets of melanoma (8 primary tumors and 9 cell lines), gastric (n=15), and breast (n=18) tumor-normal tissue pairs. A gene was considered overexpressed in a tumor sample if the expression value in the tumor sample exceeded that of the adjacent normal sample by at least two fold and the average value of all normal samples by at least three fold to account for the interindividual gene expression variations as well as to include the data from commercial normal samples and to compare those melanoma samples and cell lines that didn't have the adjacent normal sample. The expression values in tumor and normal sample groups per each tumor type were compared using the non-parametric Mann-Whitney U test for unpaired samples and Wilcoxon signed rank test for paired samples excluding the commercial normal sample, and a gene was considered significantly overexpressed if both tests gave a p-value <0.05. We found a statistically significant overexpression in cancer of two testis-selective genes – LRRC50 and ESCO1. LRRC50 was upregulated in 23% of melanoma, 30% of breast, and 60% of gastric cancer specimens, but not melanoma cell lines (Fig. 2a, b), There was no correlation of the overexpression of LRRC50 to the histological type or ER status of breast cancer, but all were early stage tumors (stages I and II), while no correlation to the tumor histological type or stage was observed for gastric cancer and melanoma samples. ESCO1 was upregulated in 53% of melanomas exceeding the level of other normal tissues by in average five times and often (in ~23% of cases) exceeding the level observed in testis (Fig. 2b).

The testis-selective genes ALLC, C11ORF20 and C16ORF82 showed upregulated expression in a few tumor samples, but didn't meet statistical significance (Fig. 2a). Neither the testis-restricted gene C21ORF66 nor the novel transcript of TEF were detected in any of the analysed tumor samples. Among the 16 ubiquitous genes four were significantly upregulated in cancer – ACTR2 in melanoma, LIG1 and NOL8 in breast cancer and melanoma, and SPARC in gastric cancer (Fig. 2a, c).

Undefined ORF antigens

The analyses scheme was similar for the undefined ORF antigens. Normal tissue analysis revealed two novel testis-selective ASVs of CFL1 and COX6B2, one novel testis-selective gene LOC284861, and two novel testis-restricted genes (LOC392843 and clone #232) (Fig. 3a, b). The expression of the putative novel genes clone #167 and clone #200 were not detected in any normal tissues by RT-PCR (Fig. 3a), but clone #200 was detected in one of the four testis samples by qPCR (Fig. 3b) also ascribing it as a novel testis-restricted gene, while clone #167 remained undetected.

Further expression analysis in melanoma and gastric tumor-normal tissue pairs showed that, although the statistical significance was not reached, the testis-selective transcripts, novel ASVs of CFL1 and COX6B2 and gene LOC284861, were upregulated in some cancer samples (Fig. 3c). The testis-restricted transcripts, LOC392843 and clone #232, were not detected in any sample. The clone #167 was first checked by conventional RT-PCR in sets of eight melanoma and gastric tumor tissues, but no positive samples were detected.

DISCUSSION

Most of the currently known clinically significant cancer antigens, such as NY-ESO1, MUC1, p53, MAGEA3 etc. that are capable to elicit T-cell immunity both spontaneously and in clinical trials, also elicit spontaneous humoral responses in cancer patients [14], hence the presence of IgG class autoantibodies is likely to reflect the lack of immune tolerance against these antigens. In this study we attempted to determine whether there are any cancer antigens of potential immunotherapeutic value among the collection of over 1300 humoral antigens identified in a previous study using phage-display SEREX ([9] and unpublished results), which is to our knowledge the largest annotated antigenic clone collection to date. Previously uncharacterised natural ORF antigens with cancer-related serum response were subjected to further ranking by the cancer autoantibody frequency, EST profile, their putative involvement in oncogenesis and structural motifs by putting a positive mark on surface localized antigens as they allow the development of additional antibody-based therapeutics and novel splice variants as they might represent novel structurally altered antigens created by deregulated alternative splicing. The top 23 rated natural ORF antigens and five cancer-related undefined ORF antigens representing novel genes were selected for further mRNA expression analysis.

LRRC50 is the top ranked antigen that was identified using a breast cancer serum for immunoscreening and further showed statistically significant overexpression and autoantibody response in this cancer type. Among the 13 breast cancer patients that had matching tumor and serum samples, four showed overexpression of LRRC50 in their tumors and two of these patients were LRRC50 antibody positive, while no corresponding antibodies were detected in sera from the 9 patients that didn't show LRRC50 overexpression in tumors. Autoantibodies associated with melanoma and gastric cancer were detected in approximately 9% of and 2% of sera, respectively. Evidence of the upregulation of LRRC50 in breast cancer and melanoma can also be found in the Oncomine database [15, 16]. Among various normal tissues it showed testis-selective expression pattern ascribing it as a novel member of the CT antigen group. There is not much known about the functions of LRRC50. Mutations in this gene have been reported to cause polycystic kidney disease in zebrafish [17], and *situs inversus* together with reduced muco-ciliary clearance of respiratory tract in humans [18], which are syndromes caused by the malfunction of motile cilia in the zebrafish kidney and embryonic nodal cells together with motile cilia-covered epithelia of the respiratory tract [19]. It is proposed that LRRC50 is necessary for the assembly of dynein-arm complexes in the cytoplasm, which is crucial for intraflagellar transport that underlies both, ciliary signaling and motile functions [18]. The non-motile primary cilium has the same structure as the motile embryonic nodal cilia with the 9+0 axoneme [20], and has recently been appreciated as a crucial nexus for proper signaling during embryogenesis and adult tissue homeostasis through the Hedgehog, PDGFRαα and Wnt pathways [21], which are also implicated in various aspects of tumorigenesis [22-24]. It will be interesting to find out if the overexpression of LRRC50 in tumor cells can participate in the signaling pathways of the primary cilium that promote the malignant phenotype considering there are no motile cilia on most cancer cells. We, however, have observed the endogenous LRRC50 on the condensed chromosomes in mitotic cells of human teratocarcinoma cell line PA1, while the interphase cells were LRRC50 negative (unpublished observations), which might indicate to another possible link to oncogenesis. Its frequency of autoantibodies and overexpression in tumor tissues is comparable to the currently widely used immunotherapy target NY-ESO1 [25] suggesting a similar immunogenicity potential.

Cancer autoantibodies recognising ESCO1 appear to be rare, but the frequency in melanomas is comparable to the well-known melanoma antigen MAGEA1 [26]. Also its

mRNA was significantly overexpressed in melanoma samples (~50%). Data deposited in the Oncomine database [16] shows substantial upregulation also in other cancer types like breast, colon, cervical, pancreatic cancer, etc. The expression pattern in normal tissues was testis-selective accounting it as a novel CT antigen. ESCO1 is an acetyltransferase important for sister chromatid cohesion during replication and DNA repair [27] by acetylating the components of the cohesion complex [28] and the repression of transcription [29]. The deregulated expression/function of ESCO1 has the potential to contribute to the genetic instability [28]. It has been proposed as the susceptibility gene of prostate cancer elevating the risk of the formation of the fusion oncogene TMPRSS2/ETS [30], and we have demonstrated here that it can be immunogenic in cancer patients. Further studies of the normal and oncogenic functions and T cell immunogenicity of these novel testis-selective antigens are warranted to evaluate their suitability as cancer immunotherapy targets.

Additionally four annotated genes with unknown functions (ALLC, C11ORF20, C16ORF82, and C21ORF66) and four novel genes (LOC284861, LOC392843, clone #232, and clone #200) with cancer-related autoantibody response were identified as testis-selective or restricted. We didn't detect the expression of any of the testis-restricted transcripts (C21ORF66, LOC392843, clone #232, and clone #200) in the analysed tumor samples suggesting that these are either very rarely expressed in tumors or in very low levels. Genes specific to germ cell have been suggested to be involved in the control and establishment of stemness [31], thereto the peptide of the novel gene clone #232 is 80% identical to the stem cell gene SLAIN1. Other approaches like mRNA expression analysis of purified tumor stem cell populations, immunohistochemistry of cancer sections or flow-cytometry of primary tumor cell cultures using antibodies against these antigens with co-staining of stem cell markers should be used to determine if they might represent genes expressed in seldom cancer stem cells that could provide additional ground for their application as immunotherapy targets. We observed the overexpression of the testis-selective genes (ALLC, C11ORF20, C16ORF82, LOC284861) in some tumour samples and these might potentially be novel CT antigens. There are 21 EST records that have yielded the model mRNA of LOC284861 and from these 17 are derived from testis or embryonic tissues, and the rest are from pooled tissues and one muscle-derived EST. The cloned antigenic peptide of LOC284861 is 124 amino acids long. No conserved domains were found, but it showed a 95% identity across the whole sequence length to a hypothetical protein from *Pan troglodytes* as well as a 79% homology to a hypothetical protein from *Pongo abelii* indicating that it might represent the natural ORF of this novel gene.

Four of the 16 analyzed cancer-related antigens representing ubiquitous genes showed statistically significant overexpression in cancer – ACTR2, SPARC, LIG1 and NOL8, and have all been reported with implications in oncogenesis. The actin remodeling protein ACTR2 determines cell shape and influences cell motility [32], and has been shown to be overexpressed in colon [33], lung [34] and other tumor types as evidenced by the Oncomine database [16]. A multifunctional matrix-associated protein SPARC influences cell shape, inhibits cell-cycle progression, and regulates the synthesis of extracellular matrix [35, 36]. It has been shown to exhibit both, tumor-promoting [37] and tumor-suppressing [38, 39] activities. LIG1 encodes DNA ligase 1 that participates in DNA replication and the base excision repair process [40], and many studies have identified SNPs in this gene associated with various cancers like glioma [41], lung cancer [42], head and neck cancer [43] etc., as well it was shown as a candidate overexpressed gene in cervical cancer [44] and other tumor types by gene array data in Oncomine [16]. NOL8 is a nucleolar protein involved in the promotion of cell growth [45], and was shown

to be overexpressed in diffuse-type gastric cancer [46]. We have shown that these genes are overexpressed in cancer, however the elevated expression level is an unlikely cause of the immune recognition as the mRNA level in the overexpressed tumors didn't exceed the level of other normal tissue types, thereto autoantibodies against ACTR2 were detected in a melanoma patient whose tumor didn't overexpress this antigen. Hence they cannot be suitable immunotherapy targets unless molecular alterations leading to cancer-specific epitope formation can be identified, and the reasons of immunogenicity of these as well as of the rest of the ubiquitous antigens remain elusive.

We also report here the identification of a large number of unnatural ORF peptide antigens derived from inserts of non-cDNA nature including intergenic regions, and we previously showed that the corresponding serum antibodies indeed recognize these unnatural peptides and not the natural ORF products of the corresponding genes [9]. Similar results have been reported by other groups using phage-display technology [47] and to a lesser extent also conventional SEREX, and likely indicate to a much broader transcriptome than previously considered supporting the recent acknowledgement of the wide variety of non-coding RNAs [48, 49]. It might be possible that the non-coding particularly the long non-coding RNA species could actually be recognized by ribosomes due to translation deregulation in cancer. However, the identified antigenic peptides most likely represent mimotopes of other antigens including proteins with altered structure or lipid and carbohydrate epitopes, the latter two representing exclusively B-cell epitopes [50]. Lately also peptide mimotopes have been suggested for cancer vaccine development as they have the immunogenic potential to elicit stronger recognition of cancer antigens of non-protein nature like the tumor-associated carbohydrate antigens already studied for the GD2 disialoganglioside and CEACAM-5 glycoprotein, as well as provide a compact immunization agent that can be coupled to additional stimulants of immune response in order to achieve an enhanced recognition of protein antigens [51].

We suspect that the large majority of the subdominant and alternative ORF peptides generated from sense direction cDNA inserts also represent mimotopes, nevertheless, there might be true novel unconventional or cryptic peptide antigens among them that are created by transcriptional mechanisms including alternative promoters, splicing defects, chromosomal translocations or translation deregulation through frameshift mutations or recognition of alternative IRES [47, 1]. Germ cells in testis possess one of the most diverse transcriptomes due to vastly rich alternative splicing. We have previously suggested that the deregulation of alternative splicing in cancer can result in the recognition of testis-restricted splice sites, leading to the production of immunogenic isoforms of otherwise tolerated proteins [52, 53]. We analyzed here the expression of antigens representing novel splice variants that showed a possible translation initiation upstream of the alternative ORF to see how often such transcripts are created in normal and cancerous tissues, and report here the identification of novel testis-selective splice variants for CFL1 and COX6B2. These transcripts were upregulated in ~20% of cancer samples, yet the statistical significance wasn't achieved. Nevertheless it suggests that the splice sites normally recognized in testis can also be recognized in cancer, but it is necessary to determine the full length structure of these transcripts and to approve the existence of the identified peptides to define these as novel cryptic peptide antigens. We recently showed the existence of a CT splice variant-specific humoral immune response for a sperm associated antigen SPAG17 [54]. We suggest that such antigens are designated to a separate category called CT-spliced antigens, as the mechanism of their production is different from the CT antigens namely splicing errors rather than transcription deregulation. However, the application of most of such transcripts in immunotherapy seems unlikely as their expression can be very heterogeneous in the same tumor due to splicing events in each

individual cell that are not inherited across cell divisions, and the antigen is produced in low level simply because the transcript represent only a fraction of the total mRNA pool of the corresponding gene.

In conclusion, we have undertaken a comprehensive gene expression analysis of the antigens eliciting humoral immune response in cancer patients that revealed several previously uncharacterized antigens with restricted expression in normal tissues. Among them 2 novel testis-selective antigens, LRRC50 and ESCO1, showed significant upregulation in cancer tissues and warrant further exploration of T cell immunogenicity. Four testis-selective antigens including three annotated genes with unknown function, ALLC, C11ORF20, C16ORF82, and one novel gene, LOC284861, were overexpressed in some tumour samples and might represent novel CT antigens. The upregulation of novel testis-selective alternative splice variants of CFL1 and COX6B2 was detected in a few tumor samples suggesting that splice sites characteristic to immunoprivileged tissues can be also recognized in cancer due to deregulated splicing and might yield immunogenic isoforms of tolerated proteins or tumor-specific cryptic peptide antigens, and we propose to designate such antigens as CT-spliced antigens.

ACKNOWLEDGEMENTS

The authors thank Dr. L. Ņikitina-Zaķe from the Genome Database of the Latvian population for the serum sample supply.

This study was supported in parts by funds from Latvian Council of Science, grant number 09.1288, Latvian State Research Program "Development of novel prophylactic, therapeutic and diagnostic tools and biomedical technologies for improvement of public health", European 6th FP project ENACT, contract number LSHC-CT-2004-503306, and European Social Fund project No. 1DP/1.1.2.1.2/09/IPIA/VIAA/004. The publishing costs are covered by ERDF project No. 2DP/2.1.1.2.0/10/APIA/VIAA/004.

REFERENCES

1. Yang F, Yang XF (2005) New concepts in tumor antigens: Their significance in future immunotherapies for tumors. Cell Mol Immunol 2 (5):331-341
2. Schietinger A, Philip M, Schreiber H (2008) Specificity in cancer immunotherapy. Semin Immunol 20 (5):276-285. doi:S1044-5323(08)00054-7/10.1016/j.smim.2008.07.001
3. Cheever MA, Allison JP, Ferris AS, Finn OJ, Hastings BM, Hecht TT, Mellman I, Prindiville SA, Viner JL, Weiner LM, Matrisian LM (2009) The prioritization of cancer antigens: A national cancer institute pilot project for the acceleration of translational research. Clin Cancer Res 15 (17):5323-5337. doi:15/17/5323/10.1158/1078-0432.CCR-09-0737
4. Finn OJ (2008) Cancer immunology. N Engl J Med 358 (25):2704-2715. doi:358/25/2704/10.1056/NEJMra072739
5. Rosenberg SA (2004) Development of effective immunotherapy for the treatment of patients with cancer. J Am Coll Surg 198 (5):685-696. doi:10.1016/j.jamcollsurg.2004.01.025/S1072-7515(04)00099-7/
6. Sahin U, Tureci O, Pfreundschuh M (1997) Serological identification of human tumor antigens. Curr Opin Immunol 9 (5):709-716. doi:S0952-7915(97)80053-2/
7. Rosenberg SA, Sherry RM, Morton KE, Scharfman WJ, Yang JC, Topalian SL, Royal RE, Kammula U, Restifo NP, Hughes MS, Schwartzentruber D, Berman DM, Schwarz SL, Ngo LT, Mavroukakis SA, White DE, Steinberg SM (2005) Tumor progression can occur despite the induction of very high levels of self/tumor antigen-specific cd8+ t cells in patients with melanoma. J Immunol 175 (9):6169-6176. doi:175/9/6169/
8. Muranski P, Restifo NP (2009) Adoptive immunotherapy of cancer using cd4(+) t cells. Curr Opin Immunol 21 (2):200-208. doi:S0952-7915(09)00019-3/10.1016/j.coi.2009.02.004
9. Kalnina Z, Silina K, Meistere I, Zayakin P, Rivosh A, Abols A, Leja M, Minenkova O, Schadendorf D, Line A (2008) Evaluation of t7 and lambda phage display systems for survey of autoantibody profiles in cancer patients. J Immunol Methods 334 (1-2):37-50. doi:S0022-1759(08)00052-5/10.1016/j.jim.2008.01.022
10. Misaki Y, Pruijn GJ, van der Kemp AW, van Venrooij WJ (1994) The 56k autoantigen is identical to human annexin xi. J Biol Chem 269 (6):4240-4246
11. Gordon T, Grove B, Loftus JC, O'Toole T, McMillan R, Lindstrom J, Ginsberg MH (1992) Molecular cloning and preliminary characterization of a novel cytoplasmic antigen recognized by myasthenia gravis sera. J Clin Invest 90 (3):992-999. doi:10.1172/JCI115976
12. web (2000) Aceview. Available via National Center for Biotechnology Information. http://www.ncbi.nlm.nih.gov/IEB/Research/Acembly/index.html. Accessed April 16, 2010
13. Hofmann O, Caballero OL, Stevenson BJ, Chen YT, Cohen T, Chua R, Maher CA, Panji S, Schaefer U, Kruger A, Lehvaslaiho M, Carninci P, Hayashizaki Y, Jongeneel CV, Simpson AJ, Old LJ, Hide W (2008) Genome-wide analysis of cancer/testis gene expression. Proc Natl Acad Sci U S A 105 (51):20422-20427. doi:0810777105/10.1073/pnas.0810777105
14. Reuschenbach M, von Knebel Doeberitz M, Wentzensen N (2009) A systematic review of humoral immune responses against tumor antigens. Cancer Immunol Immunother 58 (10):1535-1544. doi:10.1007/s00262-009-0733-4
15. Rhodes DR, Yu J, Shanker K, Deshpande N, Varambally R, Ghosh D, Barrette T, Pandey A, Chinnaiyan AM (2004) Oncomine: A cancer microarray database and integrated data-mining platform. Neoplasia 6 (1):1-6

16. web (2008) Oncomine research edition. Available via Compendia Bioscience, Inc. https://www.oncomine.org/resource/login.html. Accessed April 16, 2010
17. van Rooijen E, Giles RH, Voest EE, van Rooijen C, Schulte-Merker S, van Eeden FJ (2008) Lrrc50, a conserved ciliary protein implicated in polycystic kidney disease. J Am Soc Nephrol 19 (6):1128-1138. doi:ASN.2007080917/10.1681/ASN.2007080917
18. Loges NT, Olbrich H, Becker-Heck A, Haffner K, Heer A, Reinhard C, Schmidts M, Kispert A, Zariwala MA, Leigh MW, Knowles MR, Zentgraf H, Seithe H, Nurnberg G, Nurnberg P, Reinhardt R, Omran H (2009) Deletions and point mutations of lrrc50 cause primary ciliary dyskinesia due to dynein arm defects. Am J Hum Genet 85 (6):883-889. doi:S0002-9297(09)00473-X/10.1016/j.ajhg.2009.10.018
19. Michaud EJ, Yoder BK (2006) The primary cilium in cell signaling and cancer. Cancer Res 66 (13):6463-6467. doi:66/13/6463/10.1158/0008-5472.CAN-06-0462
20. Lancaster MA, Gleeson JG (2009) The primary cilium as a cellular signaling center: Lessons from disease. Curr Opin Genet Dev 19 (3):220-229. doi:S0959-437X(09)00086-0/10.1016/j.gde.2009.04.008
21. Veland IR, Awan A, Pedersen LB, Yoder BK, Christensen ST (2009) Primary cilia and signaling pathways in mammalian development, health and disease. Nephron Physiol 111 (3):p39-53. doi:000208212/10.1159/000208212
22. Yu J, Ustach C, Kim HR (2003) Platelet-derived growth factor signaling and human cancer. J Biochem Mol Biol 36 (1):49-59
23. MacDonald BT, Tamai K, He X (2009) Wnt/beta-catenin signaling: Components, mechanisms, and diseases. Dev Cell 17 (1):9-26. doi:S1534-5807(09)00257-3/10.1016/j.devcel.2009.06.016
24. Wong SY, Seol AD, So PL, Ermilov AN, Bichakjian CK, Epstein EH, Jr., Dlugosz AA, Reiter JF (2009) Primary cilia can both mediate and suppress hedgehog pathway-dependent tumorigenesis. Nat Med 15 (9):1055-1061. doi:nm.2011/10.1038/nm.2011
25. Gnjatic S, Nishikawa H, Jungbluth AA, Gure AO, Ritter G, Jager E, Knuth A, Chen YT, Old LJ (2006) Ny-eso-1: Review of an immunogenic tumor antigen. Adv Cancer Res 95:1-30. doi:S0065-230X(06)95001-5/10.1016/S0065-230X(06)95001-5
26. Stockert E, Jager E, Chen YT, Scanlan MJ, Gout I, Karbach J, Arand M, Knuth A, Old LJ (1998) A survey of the humoral immune response of cancer patients to a panel of human tumor antigens. J Exp Med 187 (8):1349-1354
27. Hou F, Zou H (2005) Two human orthologues of eco1/ctf7 acetyltransferases are both required for proper sister-chromatid cohesion. Mol Biol Cell 16 (8):3908-3918. doi:E04-12-1063/10.1091/mbc.E04-12-1063
28. Zhang J, Shi X, Li Y, Kim BJ, Jia J, Huang Z, Yang T, Fu X, Jung SY, Wang Y, Zhang P, Kim ST, Pan X, Qin J (2008) Acetylation of smc3 by eco1 is required for s phase sister chromatid cohesion in both human and yeast. Mol Cell 31 (1):143-151. doi:S1097-2765(08)00420-6/10.1016/j.molcel.2008.06.006
29. Choi HK, Kim BJ, Seo JH, Kang JS, Cho H, Kim ST Cohesion establishment factor, eco1 represses transcription via association with histone demethylase, lsd1. Biochem Biophys Res Commun 394 (4):1063-1068. doi:S0006-291X(10)00589-9/10.1016/j.bbrc.2010.03.125
30. Luedeke M, Linnert CM, Hofer MD, Surowy HM, Rinckleb AE, Hoegel J, Kuefer R, Rubin MA, Vogel W, Maier C (2009) Predisposition for tmprss2-erg fusion in prostate cancer by variants in DNA repair genes. Cancer Epidemiol Biomarkers Prev 18 (11):3030-3035. doi:1055-9965.EPI-09-0772/10.1158/1055-9965.EPI-09-0772
31. Simpson AJ, Caballero OL, Jungbluth A, Chen YT, Old LJ (2005) Cancer/testis antigens, gametogenesis and cancer. Nat Rev Cancer 5 (8):615-625. doi:nrc1669/10.1038/nrc1669

32. Millard TH, Sharp SJ, Machesky LM (2004) Signalling to actin assembly via the wasp (wiskott-aldrich syndrome protein)-family proteins and the arp2/3 complex. Biochem J 380 (Pt 1):1-17. doi:10.1042/BJ20040176
BJ20040176/
33. Otsubo T, Iwaya K, Mukai Y, Mizokami Y, Serizawa H, Matsuoka T, Mukai K (2004) Involvement of arp2/3 complex in the process of colorectal carcinogenesis. Mod Pathol 17 (4):461-467. doi:10.1038/modpathol.3800062
3800062/
34. Semba S, Iwaya K, Matsubayashi J, Serizawa H, Kataba H, Hirano T, Kato H, Matsuoka T, Mukai K (2006) Coexpression of actin-related protein 2 and wiskott-aldrich syndrome family verproline-homologous protein 2 in adenocarcinoma of the lung. Clin Cancer Res 12 (8):2449-2454. doi:12/8/2449/10.1158/1078-0432.CCR-05-2566
35. Bradshaw AD, Graves DC, Motamed K, Sage EH (2003) Sparc-null mice exhibit increased adiposity without significant differences in overall body weight. Proc Natl Acad Sci U S A 100 (10):6045-6050. doi:10.1073/pnas.1030790100
1030790100/
36. Bradshaw AD, Puolakkainen P, Dasgupta J, Davidson JM, Wight TN, Helene Sage E (2003) Sparc-null mice display abnormalities in the dermis characterized by decreased collagen fibril diameter and reduced tensile strength. J Invest Dermatol 120 (6):949-955. doi:12241/10.1046/j.1523-1747.2003.12241.x
37. Miyoshi K, Sato N, Ohuchida K, Mizumoto K, Tanaka M Sparc mrna expression as a prognostic marker for pancreatic adenocarcinoma patients. Anticancer Res 30 (3):867-871. doi:30/3/867/
38. Bhoopathi P, Chetty C, Gujrati M, Dinh DH, Rao JS, Lakka SS (2010) The role of mmp-9 in the anti-angiogenic effect of secreted protein acidic and rich in cysteine. Br J Cancer 102 (3):530-540. doi:6605538/10.1038/sj.bjc.6605538
39. Arnold SA, Rivera LB, Miller AF, Carbon JG, Dineen SP, Xie Y, Castrillon DH, Sage EH, Puolakkainen P, Bradshaw AD, Brekken RA Lack of host sparc enhances vascular function and tumor spread in an orthotopic murine model of pancreatic carcinoma. Dis Model Mech 3 (1-2):57-72. doi:dmm.003228/10.1242/dmm.003228
40. Tomkinson AE, Mackey ZB (1998) Structure and function of mammalian DNA ligases. Mutat Res 407 (1):1-9
41. Liu Y, Scheurer ME, El-Zein R, Cao Y, Do KA, Gilbert M, Aldape KD, Wei Q, Etzel C, Bondy ML (2009) Association and interactions between DNA repair gene polymorphisms and adult glioma. Cancer Epidemiol Biomarkers Prev 18 (1):204-214. doi:18/1/204/10.1158/1055-9965.EPI-08-0632
42. Lee YC, Morgenstern H, Greenland S, Tashkin DP, Papp J, Sinsheimer J, Cao W, Hashibe M, You NC, Mao JT, Cozen W, Mack TM, Zhang ZF (2008) A case-control study of the association of the polymorphisms and haplotypes of DNA ligase i with lung and upper-aerodigestive-tract cancers. Int J Cancer 122 (7):1630-1638. doi:10.1002/ijc.23274
43. Michiels S, Danoy P, Dessen P, Bera A, Boulet T, Bouchardy C, Lathrop M, Sarasin A, Benhamou S (2007) Polymorphism discovery in 62 DNA repair genes and haplotype associations with risks for lung and head and neck cancers. Carcinogenesis 28 (8):1731-1739. doi:bgm111/10.1093/carcin/bgm111
44. Narayan G, Bourdon V, Chaganti S, Arias-Pulido H, Nandula SV, Rao PH, Gissmann L, Durst M, Schneider A, Pothuri B, Mansukhani M, Basso K, Chaganti RS, Murty VV (2007) Gene dosage alterations revealed by cdna microarray analysis in cervical cancer: Identification of candidate amplified and overexpressed genes. Genes Chromosomes Cancer 46 (4):373-384. doi:10.1002/gcc.20418

45. Sekiguchi T, Todaka Y, Wang Y, Hirose E, Nakashima N, Nishimoto T (2004) A novel human nucleolar protein, nop132, binds to the g proteins, rrag a/c/d. J Biol Chem 279 (9):8343-8350. doi:10.1074/jbc.M305935200
M305935200/
46. Jinawath N, Furukawa Y, Nakamura Y (2004) Identification of nol8, a nucleolar protein containing an rna recognition motif (rrm), which was overexpressed in diffuse-type gastric cancer. Cancer Sci 95 (5):430-435
47. Chatterjee M, Tainsky MA (2008) Non-traditional immunogens and their application to immunotherapy. Curr Opin Mol Ther 10 (1):62-67
48. Taft RJ, Pang KC, Mercer TR, Dinger M, Mattick JS (2010) Non-coding rnas: Regulators of disease. J Pathol 220 (2):126-139. doi:10.1002/path.2638
49. Yazgan O, Krebs JE (2007) Noncoding but nonexpendable: Transcriptional regulation by large noncoding rna in eukaryotes. Biochem Cell Biol 85 (4):484-496. doi:o07-061/10.1139/o07-061
50. Meloen RH, Puijk WC, Slootstra JW (2000) Mimotopes: Realization of an unlikely concept. J Mol Recognit 13 (6):352-359. doi:10.1002/1099-1352(200011/12)13:6<352::AID-JMR509>3.0.CO;2-C
51. Knittelfelder R, Riemer AB, Jensen-Jarolim E (2009) Mimotope vaccination--from allergy to cancer. Expert Opin Biol Ther 9 (4):493-506. doi:10.1517/14712590902870386
52. Kalnina Z, Zayakin P, Silina K, Line A (2005) Alterations of pre-mrna splicing in cancer. Genes Chromosomes Cancer 42 (4):342-357. doi:10.1002/gcc.20156
53. Kalnina Z, Silina K, Line A (2008) Autoantibody profiles as biomarkers for response to therapy and early detection of cancer. Curr Cancer Ther Rev 4 (2):149-156
54. Silina K, Zayakin P, Kalnina Z, Ivanova L, Meistere I, Endzelins E, Abols A, Stengrevics A, Leja M, Ducena K, Kozirovskis V, Line A (2010) Sperm associated antigens as targets for cancer immunotherapy: expression pattern and humoral immune response in cancer patients. accepted J Immunother.

TABLES

Table 1 The classification of antigens by autoantibody frequency in each tumor type

Antigen category	Antigen subcategory	Criteria	Rating[a]	Number of antigens								
				Gastric cancer			Melanoma			Breast cancer		
				NO	ND	UO	NO	ND	UO	NO	ND	UO
Cancer-related	Significantly cancer-associated, high frequency	$p<0.05$[b],or HD=0, Ca>10%	3	2	1	4	8	2	14	4	0	14
	Putatively cancer-specific, moderate frequency	HD=0, Ca>1≤10%, $p>0.05$	2	18	1	83	19	2	108	5	0	36
	Putatively cancer-specific, rare	HD=0, Ca≤1%	1	40	2	396	53	4	366	46	7	440
	Cancer-associated	HD>0, OR≥3.5	0	8	3	23	1	4	20	15	5	74
Cancer non-related		HD>0, OR<3.5	-	62	14	381	49	9	380	60	9	323

Used abbreviations: NO – natural ORF, ND – undefined ORF, UO – unnatural ORF, HD – antibody frequency in healthy controls, Ca – antibody frequency in cancer patients, OR – odds ratio for autoantibody frequency in any type of cancer and healthy controls.

[a] The score given for each of the introduced antigen subcategories applied in further antigen prioritization scheme is indicated.

[b] p value was calculated by the χ^2 test

Table 2 Selected natural ORF antigens

Antigen	Initially identified		Serum response data										Scores				
	cDNA source	Serum	M			GC			BC			HD	Sero-logy	EST profile	Motifs	Func-tion	**Total score**[a]
			%	P-val	OR	%	P-val	OR	%	P-val	OR	%					
LRRC50	Testis	BC	8.95	**0.001**	Inf	1.73	0.293	Inf	15.38	**0.001**	Inf	0	3	3		2	**8**
SENP1	Testis	M	3.16	0.075	Inf	1.73	0.293	Inf	15.38	**0.001**	Inf	0	3	2		2	**7**
COPS4	Testis	M	1.58	0.329	Inf	0	-	NaN	0	-	NaN	0	2	2		2	**6**
RNF14	Testis+GC	GC	1.05	0.575	Inf	2.31	0.168	Inf	0	-	NaN	0	2	2		2	**6**
SLU7	Testis	M	5.26	**0.041**	8.11	1.73	0.707	2.67	15.38	**0.011**	23.69	0.65	3	2		1	**6**
SPARC	GC	GC	1.05	0.575	Inf	0.58	0.954	Inf	0	-	NaN	0	2	2		2	**6**
ACTR2	GC	GC	1.05	0.575	Inf	0	-	NaN	0	-	NaN	0	2	0	1	2	**5**
C11ORF20	Testis	M	0.53	0.572	0.81	2.89	0.286	4.45	0	0.113	0	0.65	1	3		1	**5**
C21ORF66	GC	GC	0	-	NaN	1.73	0.293	Inf	0	-	NaN	0	2	2		1	**5**
CCDC92	Testis	M	7.37	**0.008**	11.35	2.31	0.451	3.56	7.69	0.388	11.85	0.65	3	2	-1	1	**5**
RFWD2	M	M	0	0.919	0	0	0.956	0	7.69	0.388	11.85	0.65	1	2		2	**5**
ZNF282	GC	GC	2.11	0.196	Inf	2.89	0.099	Inf	0	-	NaN	0	2	2		1	**5**
TEF	Testis	M	3.16	0.075	Inf	0	-	NaN	0.79	0.129	Inf	0	2	0	1	1	**4**
C16ORF82	GC	GC	0	-	NaN	0	-	NaN	0	-	NaN	0	0	3		1	**4**
ESCO1	Testis+GC	GC	0.53	0.917	Inf	0	-	NaN	0	-	NaN	0	0	2		2	**4**
LIG1	Testis	M	1.58	0.329	Inf	0.58	0.954	Inf	7.69	0.129	Inf	0	2	0		2	**4**
NOL8	Testis	M	2.11	0.196	Inf	1.16	0.532	Inf	0	-	NaN	0	2	0		2	**4**
R3HDM2	GC	GC	2.11	0.196	Inf	0.58	0.954	Inf	0	-	NaN	0	2	0		2	**4**
RPL7A	GC	GC	2.63	0.120	Inf	1.16	0.532	Inf	0	-	NaN	0	2	0		2	**4**
RPLP1	GC	GC	0.53	0.917	Inf	1.16	0.532	Inf	0	-	NaN	0	2	0		2	**4**
RPS19	GC	GC	1.58	0.329	Inf	0	-	NaN	0	-	NaN	0	2	0		2	**4**
AIF1	GC	GC	0	-	NaN	0.58	0.954	Inf	0	-	NaN	0	0	0	1	2	**3**
ALLC	GC	GC	0.53	0.917	Inf	0.58	0.954	Inf	0	-	NaN	0	0	2		1	**3**

Used abbreviations: GC – gastric cancer, M – melanoma, BC – breast cancer, GID – gastrointestinal inflammatory disease, HD – healthy donor, % - antibody frequency, P-val – p value for frequency difference of cancer versus healthy controls calculated by the χ^2 test and bolded if <0.05, OR – odds ratio for autoantibody frequency in cancer versus healthy controls, Inf – infinity, NaN – not available.

[a] Total score obtained by summing the points given for:

Serology: 3 – significantly cancer-associated, high frequency; 2 – putatively cancer-specific, moderate frequency; 1 – cancer-associated; 0 – putatively cancer-specific, rare.

EST profile: 3 – cancer and germ cell-associated; 2 – germ cell-associated; 1 – cell line or cancer-associated; 0 – no association to cancer or germ cells.

Motifs: 1 – cell surface localization or novel alternative splice variant; -1 – coiled coil domains.

Function: 2 – potentially oncogenic, with published circumstantial evidence; 1 – unknown relation to oncogenesis; 0 – known function, no relation to oncogenesis

Table 3 Selected undefined ORF antigens

	Antigen name	Initially identified		Serum response data										Description; accession number
		cDNA source	Serum	M %	M p-val	M OR	GC %	GC p-val	GC OR	BC %	BC p-val	BC OR	HD %	
Novel splice variants	CFL1	Testis	M	1.05	0.575	Inf	0	-	NaN	0	-	NaN	0	Novel alternative 5'splice site of intron 1 elongating exon 1 in Aceview transcript c; peptide of a putative subdominant ORF; HM370399.
	COX6B2	Testis	M	4.74	**0.019**	Inf	2.31	0.168	Inf	0	-	NaN	0	Putative novel 5'splice site of intron 6 elongating exon 7 of Aceview transcript j; peptide of a putative subdominant ORF; HM370404.
Novel genes	LOC284861	Testis+GC	GC	0	-	NaN	0	-	NaN	0	-	NaN	0	Hypothetical gene supported by BC039313, fragment of Aceview transcript b-u; potential natural ORF peptide; BC039313
	LOC392843	Testis	M	0	-	NaN	0.58	0.954	Inf	0	-	NaN	0	Hypothetical gene model mRNA XR_041603.2, contains novel exon in intron 10 of this model mRNA; frame 3 peptide of the predicted protein; HM370405.
	Clone #232	Testis	PC	4.21	0.095	6.48	6.36	**0.019**	9.79	0	0.113	0	0.65	Conventionally spliced -str sequence of chr 5, this region contains features on the + strand: PCDHB19P and predicted PCDHB15 precursor (hmm864524), UniGene clusters Hs.570898, Hs.130757, Hs.287793, Hs.570898; 14 aa, contains STOP, putative natural ORF peptide, 80% identity to stem cell gene SLAIN1; HO000321.
	Clone #200	Testis	M	0	-	NaN	1.73	0.293	Inf	0	-	NaN	0	Conventionally spliced -str sequence of chr 11, this region contains features on the + strand: predicted gene hmm1684754, UniGene cluster Hs.418450; 66 aa, no STOP, potential natural ORF peptide, no homology to any known proteins; HO000322.
	Clone #167	Testis	M	9.47	**0.013**	4.8	6.94	0.077	3.51	15.38	0.090	7.8	1.94	Conventionally spliced -str sequence of chr 3, this region contains features on the - strand: predicted alternative isoform of RPL32 (hmm6100903), and on the +str: hypothetical gene hmm1294904, UniGene clusters Hs.265174 and Hs.343664; 55 aa, contains STOP, potential natural ORF, 50% homology to predicted human protein hCG1999481; HO000323.

Used abbreviations: GC – gastric cancer, M – melanoma, BC – breast cancer, HD – healthy donor, PC – prostate cancer, % - antibody frequency, P-val – p value for frequency difference of cancer versus healthy controls calculated by the χ^2 test and bolded if <0.05, OR – odds ratio for autoantibody frequency in cancer versus healthy controls, Inf – infinity, NaN – not available, chr – chromosome, +str – plus strand, -str – minus strand, STOP – translation termination codon

FIGURE LEGENDS

Fig. 1 The expression analyses of natural ORF antigens in various normal tissues. **A** RT-PCR was performed on 11 various normal tissues. The shading intensity relatively corresponds to the intensity of the gel band. White – not detected, light gray – weak intensity band, medium gray – medium intensity band, black – strong intensity band. **B** qPCR was performed on 14 various normal tissues to verify testis-associated expression pattern of genes detected in RT-PCR. Y axis represents percentage of the testis level and is indicated in logarithmic scale. **C** qPCR results of testis-selective genes LRRC50 and ESCO1 relatively to the testis sample in linear scale

Fig. 2 The expression analyses of natural ORF antigens in tumor tissues. qPCR was performed in sets of melanoma (M) (8 primary tumors, 9 cell lines and 5 unpaired normal skin samples), gastric (G) (15), breast (B) (18) tumor-normal tissue pairs. **A** Volcano plot of qPCR expression data of all analyzed genes and tumor sample sets. Y axis represents minus log10 of normally approximated p value obtained by non-parametric Mann-Whitney test. Dashed line crossing Y axis represents the 0.05 p-value. X axis represents log2 of obtained fold changes between average tumor and normal sample sets. Dashed lines crossing X axis represent up and downregulation by 2 fold. **B** The expression values of novel CT antigens are indicated. The median values with interquartile ranges are displayed as box plots (dark grey – tumour tissues, light grey – cell lines, white – normal tissues) with Tukey whiskers overlaid with individual sample values as a scatter dot plot. Normally approximated p-values obtained by non-parametric Mann-Whitney test are indicated if <0.05. Numbers below the tumor type designation indicate the overexpressed tumor samples out of all analysed. **C** The expression values of potential overexpressed antigens are indicated. The median values with interquartile ranges are displayed as box plots (dark grey – tumour tissues, light grey – cell lines, whilte – normal tissues) with Tukey whiskers overlaid with individual sample values as a scatter dot plot. Normally approximated p-values obtained by non-parametric Mann-Whitney test are indicated if <0.05. Numbers below the tumor type designation indicate the overexpressed tumor samples out of all analysed

Fig. 3 The expression analyses of undefined ORF antigens. **A** RT-PCR was performed on 11 various normal tissues. The darkness of the shading relatively corresponds to the intensity of the gel band. White – not detected, light gray – weak intensity band, medium gray – medium intensity band, black – strong intensity band. **B** qPCR was performed on 14 various normal tissues to verify testis-associated expression pattern of genes detected in RT-PCR. Y axis represents percentage of the testis level and is indicated in logarithmic scale. **C** qPCR was performed in sets of melanoma (M) (8 primary tumors, 9 cell lines and 5 unpaired normal skin samples), and gastric (G) (8) tumor-normal tissue pairs. The individual sample values are shown as a scatter dot plot connecting paired tumor (dark grey) and normal (white) samples as well as cell lines (light grey) in tumor types that exhibited specimens with upregulated expression

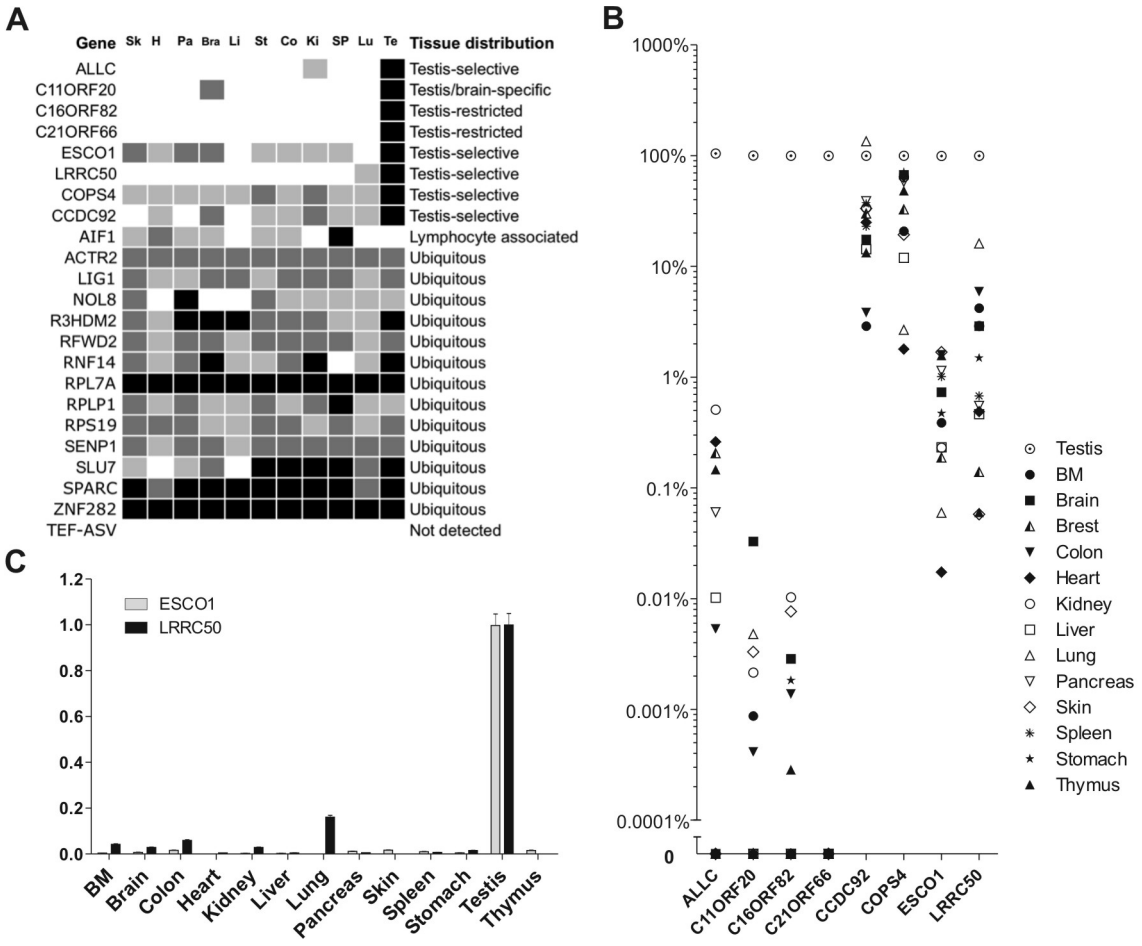

A
Gene
Sk H Pa Bra Li St Co Ki SP Lu Te
Tissue distribution
ALLC Testis-selective
C11ORF20 Testis/brain-specific
C16ORF82 Testis-restricted
C21ORF66 Testis-restricted
ESCO1 Testis-selective
LRRC50 Testis-selective
COPS4 Testis-selective
CCDC92 Testis-selective
AIF1 Lymphocyte associated
ACTR2 Ubiquitous
LIG1 Ubiquitous
NOL8 Ubiquitous
R3HDM2 Ubiquitous
RFWD2 Ubiquitous
RNF14 Ubiquitous
RPL7A Ubiquitous
RPLP1 Ubiquitous
RPS19 Ubiquitous
SENP1 Ubiquitous
SLU7 Ubiquitous
SPARC Ubiquitous
ZNF282 Ubiquitous
TEF-ASV Not detected
B
1000%
100%
10%
1%
0.1%
0.01%
0.001%
0.0001%
0
ALLC
C11ORF20
C16ORF82
C21ORF66
CCDC92
COPS4
ESCO1
LRRC50
Testis
BM
Brain
Brest
Colon
Heart
Kidney
Liver
Lung
Pancreas
Skin
Spleen
Stomach
Thymus
C
ESCO1
LRRC50
1.2
1.0
0.8
0.6
0.4
0.2
0.0
BM
Brain
Colon
Heart
Kidney
Liver
Lung
Pancreas
Skin
Spleen
Stomach
Testis
Thymus

Figure 1

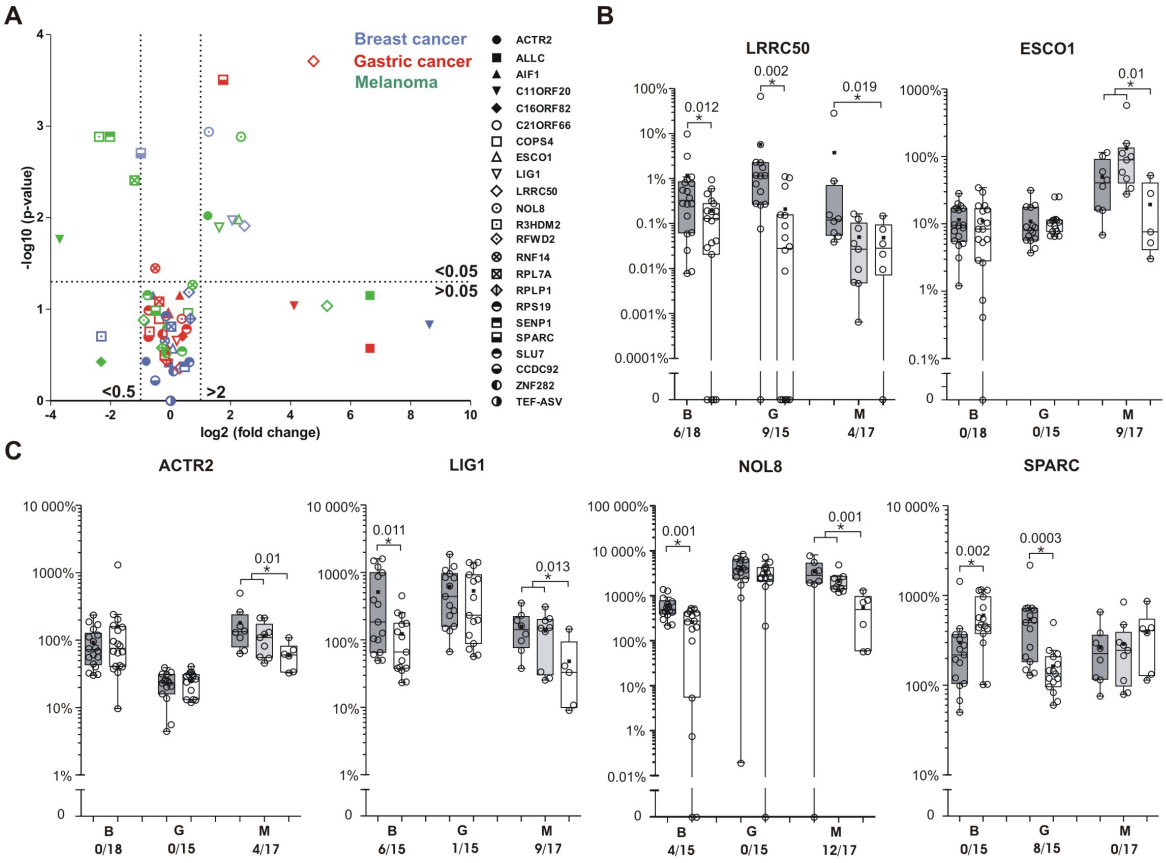

A
Breast cancer
Gastric cancer
Melanoma
ACTR2
ALLC
AIF1
C11ORF20
C16ORF82
C21ORF66
COPS4
ESCO1
LIG1
LRRC50
NOL8
R3HDM2
RFWD2
RNF14
RPL7A
RPLP1
RPS19
SENP1
SPARC
SLU7
CCDC92
ZNF282
TEF-ASV
-log10 (p-value)
log2 (fold change)
<0.05
>0.05
<0.5
>2
B
LRRC50
ESCO1
C
ACTR2
LIG1
NOL8
SPARC

Figure 2

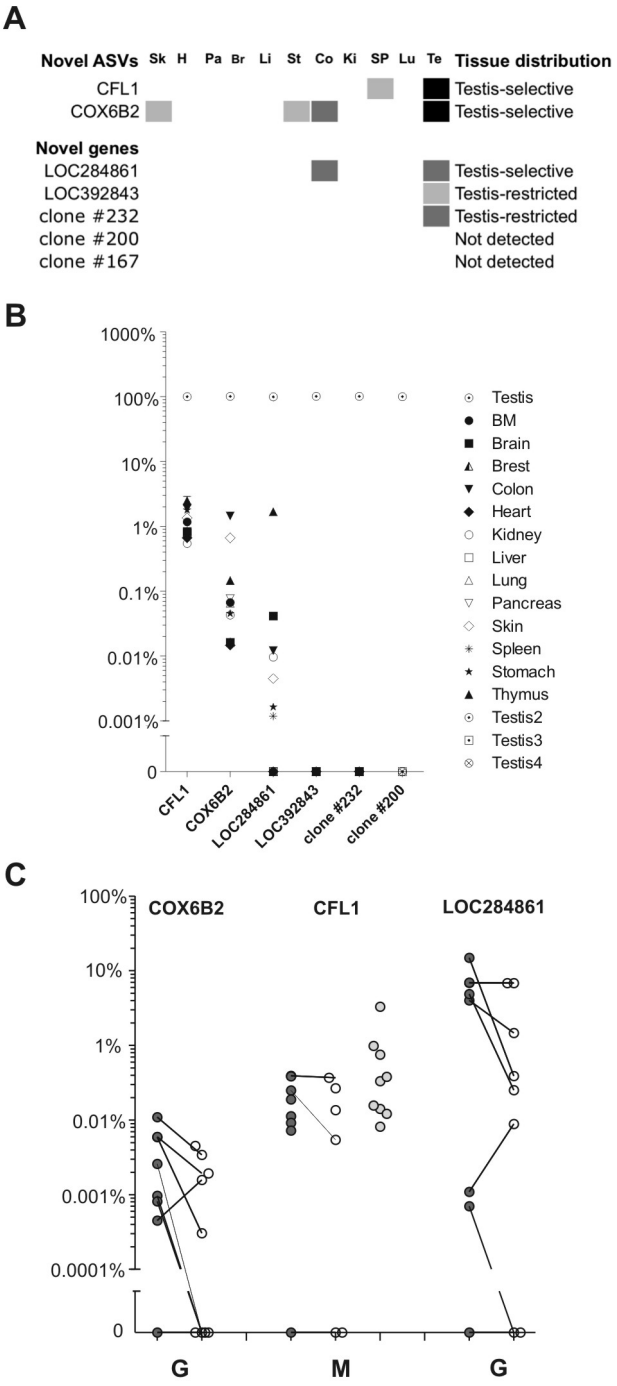

A
Novel ASVs
Sk H Pa Br Li St Co Ki SP Lu Te
Tissue distribution
CFL1
COX6B2
Testis-selective
Testis-selective
Novel genes
LOC284861
LOC392843
clone #232
clone #200
clone #167
Testis-selective
Testis-restricted
Testis-restricted
Not detected
Not detected
B
1000%
100%
10%
1%
0.1%
0.01%
0.001%
0
CFL1
COX6B2
LOC284861
LOC392843
clone #232
clone #200
Testis
BM
Brain
Brest
Colon
Heart
Kidney
Liver
Lung
Pancreas
Skin
Spleen
Stomach
Thymus
Testis2
Testis3
Testis4
C
100%
10%
1%
0.01%
0.001%
0.0001%
0
COX6B2
CFL1
LOC284861
G
M
G

Figure 3

Table S1. Primer sequences used in antigen mRNA expression analysis.

Analysis	Gene	Forward primer 5'-3'	Reverse primer 5'-3'
Reference genes	PGK1	CTTAAGGTGCTCAACAACATGG	ACAGGCAAGGTAATCTTCACAC
	POLR2A	GGGTCATCTTCCCAACTGGAG	CACCAGCTTCTTGCTCAATTCC
	TBP	CCACTCACAGACTCTCACAAC	CTGCGGTACAATCCCAGAAC
	ACTB	AATCTCATCTTGTTTTCTGCGC	AGTGTGACGTGGACATCCG
	GAPDH	GGGTCTTACTCCTTGGAGGC	GTCATCCCTGAGCTAGACGG
	TUBA3-R	TATGGCAAGAAGTCCAAGCTG	TACCATGAAGGCACAATCAGAG
	YWHAZ	CCTGCATGAAGTCTGTAACTGAG	GACCTACGGGCTCCTACAAC
	ACTR2	ATTGCTGGGAGGGATATAAC	TTTCTAAGGCCAGTTTCTGC
Natural ORF antigens	ACTR2	ATTGCTGGGAGGGATATAAC	TTTCTAAGGCCAGTTTCTGC
	AIF1	CATGTCCCTGAAACGAATGC	TAGGATGGCAGATCTCTTG
	ALLC	ACTGCTTCCAGTGACCAAGT	GTGTGTTAAGGGTTTGCTTTG
	C11ORF20	AGCCTTCGAAAGTGTCGCTC	CCTCGCAGACCTCGGACAG
	C1ORF82	GAATGAAGCAGGAAGGAGAG	GCTGCACGCTGGAAACAT
	C21ORF66	CTGGAGAACAGGATGAAGAG	CTATAAGGAATGCCATAGG
	CCDC92	AGCTGGAAGCCCAACTGAAAG	CAGGTAGGCGATGGTGCTG
	COPS4	CTGGAGCAGAACAACTAGAAG	CAAATGAAATGACTCTAGGCTG
	ESCO2	CTGCAATCTGCGGGATCAGTC	GTACTGTGTTGCAAACAGC
	LIG2	AGGAGTGGAATGGAGTGGTG	TGGGAGAGGTGTCAGAGAGG
	LRRC50	CAAAGAGATGCTGCACCACTC	CTATGATGCTTTCGGTGCTGG
	NOL8	TTTACCACAGTCCTCATTCC	ACAAACAGAGGATCAGTAGC
	R3HDM2	GAGCAGCACAGACAGCGAAC	GACCTGGTTGCACACTTCTG
	RFWD2	AGTCTCTAGTATTGAATTTGACC	CATTCTCAGGGTAATGAATATCC
	RNF14	GGAACTCCCATAGAGAAATTAG	CCCAAATATCGTCGTCAACATC
	RPL7A	GAATTTTGGCATTGGACAGG	CTGGGTGAACTGGTTAATCG
	RPLP1	CCAACGTCAACATTGGGAG	GGTTTAGTCAAAAAGACCAAA
	RPS19	GGAAAAGGACCAAGATGGC	GTTCTAATGCTTCTTGTTGGC
	SENP1	CAGTGAAGCATTTCGCCTGAC	AACGTTTCACTGCCTGATAACC
	SLU7	TCCAGATGAACATGTCCAGCC	TCTCCCCACTGGTGTTTTGG
	SPARC	GCAGAGGTGACTGAGGTATC	CCGTGTTTGCAGTGGTGGTT
	TEF-ASV	AGGCGCGCCTCGAATCTTC	AGCAAACTTGTGCTTCCGAGG
	ZNF282	GAAGAACCTTGTGTGTGGGAG	CAGAAATTGGGGAGTCGGT
Undefined ORF antigens	CFL1	GGTATGCTCTGCAAGGCCTC	CTTCTTGCGCTTCTTCACCTC
	Clone #167	TGGCTTTGATGTTGAGTCTCC	AAAATACGAAAACCAGTCAGG
	Clone #200	CCAAATAGCTGGATTACAGAA	GTGGTTGATGTCATTGTTTG
	Clone #232	GAGGCGAGGGGCAGAATC	GCACTTAATGAACGAAGGGGAG
	COX6B2	CCAGAGCCTTGACGAATGCAC	CAGGAGGCTGCTGTGGCTG
	EVI5L	AGCTACAGGAGCAGCAGG	GCGCTTGTCTCTGGGTTC
	LO392843	CCCCGCTGAGGTCTTTGTG	CCAAATGTATCCCTATGATCTG
	LOC284861	TTCTCCAGTGTCAATGGCTCA	ACACTGTACAAGACATCTGTAA

BASIC STUDY

Sperm-associated Antigens as Targets for Cancer Immunotherapy: Expression Pattern and Humoral Immune Response in Cancer Patients

Karīna Siliņa, Pavel Zayakin,* Zane Kalniņa,* Lāsma Ivanova,* Irēna Meistere,* Edgars Endzeliņš,* Artūrs Ābols,* Aivars Stengrēvics,† Mārcis Leja,† Kristīne Ducena,‡ Viktors Kozirovskis,§ and Aija Linē**

Summary: The identification of novel cancer-related and immunogenic proteins is still a challenge to be faced to improve antigen-specific tumor immunotherapy. The category of so-called cancer-testis (CT) antigens is one of the most perspective groups of proteins for anticancer immune response activation as normally they are expressed in immunoprivileged tissues and are immunogenic if aberrantly generated in tumors. The heterogeneous group of proteins called sperm-associated antigens (SPAG) might encompass novel CT antigens owing to their common expression in male germ cells, their ability to elicit immune response underlying infertility, and lately proposed oncogenic properties. We carried out a comprehensive analysis of the expression pattern in various normal and cancerous tissues and assessed the frequency of spontaneous humoral immune response against members of the SPAG group in cancer patients using phage-displayed antigen microarrays. Our results show that out of 15 analyzed SPAG genes only SPAG1, SPAG6, SPAG8, SPAG15, and SPAG17 are predominantly expressed in testis, whereas the others are ubiquitously expressed with only a testis-associated alternative splice variant of SPAG16. mRNA expression of SPAG1, SPAG6, and alternative splice variants of SPAG8, SPAG16, and SPAG17 was elevated in various tumors with frequencies ranging from approximately 10% to 70%. The upregulation of SPAG6 in lung and breast cancer was confirmed by immunohistochemical analysis of tumor and normal tissue microarrays. Cancer-associated spontaneous humoral immune response was detected against SPAG1, SPAG6, SPAG8, and a novel testis-specific splice variant of SPAG17 and ascribe these as novel CT antigens that potentially are applicable as immunotherapeutic targets and serologic biomarkers.

Received for publication July 9, 2010; accepted August 23, 2010.
From the *Latvian Biomedical Research and Study Centre; †Riga Eastern Clinical University Hospital, Oncology Centre; ‡Pauls Stradins Clinical University Hospital, Department of Endocrinology; and §Pauls Stradins Clinical University Hospital, Clinic of Oncology, Riga, Latvia.
Supported in parts by funds from Latvian Council of Science, grant numbers 09.1288 and 09.1310, Latvian State Research Program "Development of novel drugs: construction, transport forms and mechanisms of action", European 6th FP project ENACT, contract number LSHC-CT-2004-503306, and European Social Fund project number 1DP/1.1.2.1.2/09/IPIA/VIAA/004. The publishing costs are covered by ERDF project number 2DP/2.1.1.2.0/10/APIA/VIAA/004.
All authors have declared that there are no financial conflicts of interest in regards to this work.
Reprints: Karīna Siliņa, Latvian Biomedical Research and Study Centre, Ratsupites 1, Riga, LV-1067, Latvia (e-mail: karina@biomed.lu.lv).
Supplemental Digital Content is available for this article. Direct URL citations appear in the printed text and are provided in the HTML and PDF versions of this article on the Journal's website, www.immunotherapy-journal.com.

Key Words: sperm-associated antigen, CT antigen, CT-spliced antigen, antigen microarray, cancer immunotherapy

(*J Immunother* 2011;34:28–44)

Tumor antigen-specific immunotherapy holds a great potential to eradicate cancer in a highly specific and nontoxic manner, and allows direct monitoring of immune response, which is crucial for understanding immunologic mechanisms that underlie tumor regression in vivo.[1] However, the considerably low efficacy in clinical trials has hampered the translation of such therapy approaches into clinical practice and is caused by the selection of tumor escape variants through inefficient antigen presentation, the loss of antigen expression and the immunosuppressive activities of tumor, its stroma and/or innate, and adaptive tolerogenic immune cells.[2,3] Attempts have been made to overcome the above hurdles by combining antigen targeting with blocking of tolerance, adding adjuvants, altering ways of antigen delivery, TCR-based gene therapy etc., but there is still a risk of tumor escape through down-regulation of target antigen expression in case the antigen is not necessary for tumor survival and if polyclonal T-cell activation by epitope spreading was not achieved.[4] Hence, the need to develop polyvalent antigen targeting approaches as one of the means to avoid the selection of tumor escape variants, and to cover the heterogeneity of tumors, and the need to determine the protective antigens in tumors that do not express the currently exploited targets like WT1, MUC1, Her-2/neu, NY-ESO-1, and others[5] puts the discovery of novel tumor antigens in the frontline of successful antigen-specific immunotherapy development.

Cancer-testis (CT) antigens are perspective candidates for cancer immunotherapy, as naturally they are expressed in immunoprivileged tissues, but are detected in various neoplastic lesions, and they can induce spontaneous immune responses when aberrantly produced in tumors as there is no or weak central tolerance against proteins restricted to immunoprivileged sites. Emerging evidence suggests that CT antigens may possess functions related to stemness owing to their expression during germ cell and embryonic development, hence, providing space for crucial oncogenic properties in cancer cells.[6] One of the most successful antigen-specific immunotherapy targets to date is the CT antigen NY-ESO-1[5]; its targeting has shown to underlie tumor regression and induce integrated immune system activation involving both, cellular and humoral responses.[7–10]

During the last 2 decades, a group of proteins called sperm-associated antigens (SPAG) has grown to reach 15 members—SPAG1,[11] SPAG2/UAP1,[12] SPAG4,[13] SPAG5,[14–16] SPAG6,[17] SPAG7,[18] SPAG8,[19] SPAG9,[20] SPAG10/MFGE8,[21] SPAG11B,[22] SPAG12/NHP2L1,[23] SPAG13/SSFA2,[24] SPAG15/SPAM1,[25] SPAG16,[26] and SPAG17.[27] For further simplicity in this report, we use the names of the SPAG group even if they differ from the official gene name. These are functionally and evolutionary distinct proteins with a common expression in testis or sperm, and humoral immune response against some have been shown to underlie infertility. Several of them have been candidates for treating infertility[23] and for the development of immunocontraception.[28] Lately also, a role in tumorigenesis has been ascribed to several SPAG proteins, for example, SPAG1 has been shown to be a progression marker of pancreatic cancer and promote cell motility.[29] Decreased expression of SPAG5, a mitotic spindle protein, also known as astrin, is associated with good prognosis for estrogen receptor positive breast cancer patients as determined by cDNA microarray data analysis.[30] In our earlier study of large-scale identification of humoral tumor antigens, we found SPAG8 to react with sera from melanoma patients,[31] and it has been shown to be overexpressed in cervical carcinoma.[32] SPAG9 or JLP has been reported to facilitate migration and invasiveness of renal cell carcinoma[33] and proposed as an early marker for breast[34] and cervical[35] cancers. Besides, it was shown to be an AML[36] and epithelial ovarian cancer-specific antigen and was suggested for cancer immunotherapy applications,[37] and together with SPAG4 are designated as CT genes by the CT Database.[38] The vastly studied SPAG10 also known as lactadherin or the milk fat globule protein MFGE8 was identified as a breast cancer-specific antigen and was considered as a perspective serologic diagnostic marker and a radioimunotherapy target.[39,40] Elevated level of SPAG15 or testis-specific hyaluronidase PH-20 might contribute to the metastatic potential of breast and laryngeal cancer.[41–45] Besides, several SPAG proteins have been related in one way or another to the primary cilium, which is a crucial regulator of Hedgehog, PDGFaa, and WNT signaling pathways that can underlie various aspects of tumorigenesis.[46,47] Noteworthy, several of the SPAG proteins have been shown to be located on the sperm surface, which mediates infertility in patients with corresponding sperm agglutinating antibodies.[11,19] If SPAG proteins were also present on the tumor cell surface, novel-specific antibody therapeutics could be developed. Hence, SPAGs might represent a novel group of CT antigens with functional implications in cancer formation and might have a potential to be novel targets for tumor immunotherapy. However, many SPAG proteins have been analyzed in a narrow focused way concerning infertility. The aim of this study was to gain view on the expression profile and the frequency of spontaneous humoral immune responses against SPAG proteins by using phage-displayed antigen microarrays to determine if any member of the SPAG group might fulfill the requirements for a novel tumor immunotherapy target.

MATERIALS AND METHODS

Patient Material, Cell Lines

Tumor and adjacent normal tissue specimens were obtained from operation material of gastric, colon, melanoma, and breast cancer patients undergoing surgery in Latvian Oncology Centre and stored in RNALater (Applied Biosystems). Lung tumor specimens were obtained during standard diagnostic bronchoscopy procedure at Pauls Stradins Clinical University Hospital and stored in RNALater. All patients have signed an informed consent, and the study has been approved by the Central Commission of Medical Ethics of Latvia.

Sera were collected from the same patients whose tissue specimens were collected. Additional serum samples of melanoma, lymphocytic leukemia, gastric, lung, colon, breast, and thyroid cancer patients and healthy individuals with no known history of cancer, infertility, and autoimmune disorders were provided by the Genome Database of the Latvian population and sera from melanoma, gastric, and prostate cancer patients were kindly provided by the Skin Cancer Unit in German Cancer Research Center, the Clinic of Gastroenterology, Hepatology and Infectious Diseases, Otto-von-Guericke University Magdeburg, Norwegian Radium Hospital, and Onyvax Vaccine Therapies Ltd, UK.

RNA Extraction, cDNA Synthesis

Bead-based tissue homogenization was done by using the FastPrep-24 instrument and Lysing Matrix A (MP Biomedicals) in 1 mL of TRIsolution (Applied Biosystems) followed by the extraction of total RNA according to manufacturer's protocol. RNA from melanoma cell lines was kindly provided by Skin Cancer Unit in German Cancer Research Center. RNA of various normal tissues was purchased from Applied Bisystems, USA and Biocat, Germany. Total RNA was treated with DNase I using DNA-*free* DNase treatment and removal reagents (Applied Biosystems). cDNA was synthesized by random hexamer priming from 2 μg of total RNA by using RevertAidTM First StrandcDNA Synthesis Kit (Fermentas, Lithuania) according to manufacturer's instructions.

mRNA Expression Analysis

Qualitative RT-PCR reaction mixtures contained 1 × reaction buffer, 2.5 mM $MgCl_2$, 0.1 μM primers (Table 1), 0.75 U Taq DNA polymerase (Fermentas, Lithuania), and 1/60th of the prepared cDNA. Amplification of all target and reference genes was done at the same cycling conditions (45 s at 94°C, 30 s at 58°C, 45 s at 72°C), except for the number of cycles that was determined individually according to mRNA abundance (Table 1).

Quantitative RT-PCR (qPCR) reactions were done using 1/60th of cDNA reaction mixture, ABSolute Blue SYBR green Low ROX (Thermo Scientific) and ABI7500 sequence detection system (Applied Biosystems). Appropriate primer concentrations were established by serial dilution curves to ensure amplification efficiency over 95% (Table 1). To normalize the expression data a normalization factor was calculated for each cDNA from the expression values of the 3 most stable reference genes (ACTB, POLR2A, TUB3A) determined among 7 most often used housekeeping genes (GAPDH, ACTB, POLR2A, TUB3A, TBP, YWHAZ, PGK1) by using geNorm software.[48] All reactions were carried out in duplicates.

The statistical analysis of the expression data from tumor and normal samples was done using the nonparametric Mann-Whitney *U* test.

TABLE 1. Primers and Cycling Conditions Used in mRNA Expression Analyses

Gene	Forward Primer, 5′-3′	Reverse Primer, 5′-3′	RT-PCR Cycles	qPCR Ratio F/R (nM)
SPAG1	GAAAAGCATCTTCAAGCCTTGG	GGAGGTCAAGCACCAAGTTTG	35	100/100
SPAG4	TGGGTCTCCAGTAGTCTCTGA	TCCTCTGCACGACCAGTCG	35	—
SPAG5	CGCAGAGCAGGTTCAAACAC	GGAGAGGCACTTGAATGGGA	38	—
SPAG6	AGCAATGGCAGTCATCATTTC	GGATGAATGGTCGGGAACTT	35	50/100
SPAG8	CAAGCATGCAGGATGGCTCT	ATGGCTTCACGCTTCCCTCG	35	100/100
SPAG8-e2L	CTACAACTGGGAGGAAGAGAG	GTGGCTGGTACGAGTCTTTC	35	100/100
SPAG9	AGACCCGAGTGGAATCTTTAG	GTTGATCACTCCCTGAGAGC	35	—
SPAG9-C	CTCATACCAGCCTGAAGGTC	CCATCGGGTCCTTTGATCTT	35	100/50
SPAG11B	CACCCAGCCTCACTCCATC	CACTTTGCCCTGGAGAATGG	35	—
SPAG12	CAAGAAGCTACTGGACCTCG	GATGCACAAACACGCGAGG	35	—
SPAG13	GCACCACTGACAATACCATCC	CACGACTATCAACACTGTCACT	35	—
SPAG15	GTTGCTCTGGGTGCTTCTG	GGTCCTCGTTCCTCACACA	35	100/50
SPAG16	CGGAAAACAGTTCTTCCTTC	AGACTGAAAGCAATCAAGAG	35	100/100
SPAG16-L	CTGTCTATATGGGATGCAAGAAC	GACCGTTACTCCACTTAAAACTA	35	50/100
SPAG17	AACAGAAATCCTCAAGTGTGC	TGTGTTCACTTTTCCTCCAAC	35	100/100
SPAG17-A	GGAACATTGCTCTCCACTCC	GCTAATCGTCTTCTCCTCGC	35	50/100
SPAG17-A1	CAACATGAGTCTCTGGGTAA	GCTAATCGTCTTCTCCTCGC	38	100/100
SSX2	GCTCAAATACCAGAGAAGATCC	GTGGCCTTGAAACCTAGTTTAG	35	100/50
ACTB	AATCTCATCTTGTTTTCTGCGC	AGTGTGACGTGGACATCCG	25	100/100
PolR2A	GGGTCATCTTCCCAACTGGAG	CACCAGCTTCTTGCTCAATTCC	—	100/100
TUBA3	TATGGCAAGAAGTCCAAGCTG	TACCATGAAGGCACAATCAGAG	—	100/100

Immunohistochemical Analysis of Tissue Microarrays

Paraffin-embedded tissue microarrays (TMA) of various normal tissues (duplicated 45 tissues, A700 (III), AccuMax Array triplicated 33 tissues, FDA994, US Biomax, Inc.) and melanoma, gastric (ME481t and ST805t, US Biomax, Inc.), lung, and breast [A716 (III) and A712 (13), Accu Max Array] cancers were used for standard immunohistochemistry analysis including melanin bleaching procedure (Melanin Bleach Kit, Polysciences, Inc.) for normal tissue and melanoma TMAs according to manufacturers protocol. In brief, after standard deparaffination and hydration, TMAs were incubated in melanin bleaching solutions, rinsed, and continued with peroxidase quenching. Antigen retrieval was done by heating the TMA in 0.01-M citric acid, pH 6.0 (SIGMA-Aldrich, Germany) for 15 minutes. Commercial primary antibodies for SPAG6 (mouse monoclonal antibody, SantaCruz Inc), and for SPAG8 (rabbit polyclonal antibody, Proteintech Group, Inc.) were used at the dilution 1:50 and incubated overnight at +4°C. Secondary antibodies (HRP conjugated antimouse IgG antibody and antirabbit IgG antibody, SIGMA-Aldrich, Germany) were used at the dilution 1:100 for 1 hour at +37°C. DAB (SIGMA-Aldrich, Germany) was used for color development and counterstained with hematoxylin (SIGMA-Aldrich, Germany).

Analysis of Autoantibody Responses

The antigenic regions of SPAG proteins were predicted using algorithms developed by Welling et al[49] for SPAG1, SPAG6, SPAG8, SPAG9, SPAG16, and SPAG17, which, together with full CDS regions of well-known CT antigens NY-ESO-1 (CTAG1B), MAGEA1, HORMAD1, and SSX2 were amplified using high-fidelity PCR enzyme mix (Fermentas, Lithuania) from testis cDNA. Several different transcripts were obtained for SPAG6 (designated as SPAG6-A, B, and A1), SPAG9 (designated as SPAG9-A and C), and SPAG17 (designated as SPAG17-A and A1). All fragments were cloned into T7Select 10-3b Phage Display vector (Novagen) (Table, Supplemental Digital Content 1, http://links.lww.com/JIT/A84 for cloned antigens with corresponding amino acid positions). StrepII tag (Trp-Ser-His-Pro-Gln-Phe-Glu-Lys) was inserted into HindIII and NotI sites located downstream the cDNA cloning site of the vector DNA to monitor the copy number of the recombinant proteins on each phage. Obtained recombinant phages together with 9 nonrecombinant control phages were amplified, purified, and printed in triplicate on nitrocellulose slides (Whatman, GE Healthcare) to create an antigen microarray, which was screened with sera from 39 breast, 24 lung, 33 colon, 28 thyroid, 172 gastric (stage I—33, II—20, III—31, IV—37, not determined—53), 52 prostate cancer, 163 melanoma (stage I—22, II—23, III—22, IV—22, not determined—74), and 28 lymphocytic leukemia patients and 127 patients of gastrointestinal inflammatory diseases (gastric ulcer—17, duodenum ulcer—20, gastritis and duodenitis—10, acute hemorrhagic gastritis—13, chronic atrophic gastritis—48, dyspepsia—11, Crohn disease—8) and 147 healthy donors. The production and processing of antigen microarrays were done as described earlier.[31] In brief, the microarray slides were incubated with patients' sera and mouse anti-T7 phage tail antibody (Novagen). The serum reactivity was detected by Cy5 conjugated goat anti-human IgG antibody (Jackson ImmunoResearch) and the ratio against the total amount of printed phage, which was detected by Cy3 conjugated goat anti-mouse IgG antibody (Jackson ImmunoResearch) was calculated. A 2-step normalization strategy was used for the fluorescent signal ratios to eliminate variations introduced by custom production of microarrays and variable background intensities of different sera. At first, the values in each slide (each serum) were normalized by the median of all printed spots for each fluorescent channel separately. Next, the Cy5 and Cy3 signal intensities for each spot were divided by the median value of that spot within the print lot. The threshold value of a specific antibody response for each antigen was defined as 4 standard deviations above the

TABLE 2. Summary of Published Literature and Open Source Expression Data

SPAG Gene	Official Name	Evidence of Functions	Immunogencity	Expression Evidence	
				Normal Tissues*	Tumors†
SPAG1	SPAG1	Signaling from G protein-coupled receptors during spermatogenesis and fertilization[59,60] through PKC dependent EKR activation[61]; mediates maternal mtDNA inheritance,[62,63] promotes cell motility[29]	Infertility related[11,59]	Testis-selective (also in gastrointestinal tract, pancreas, tonsils, lung, skin, liver, kidney)[11,29,59,64]; colon, appendix[57]; tracheal, bronchial and nasal epithelial cells, colon[55]	Upregulated in pancreatic cancer,[29] seminomas,[65] renal, breast and other cancers,[53] colon cancer[55]
SPAG2	UAP1	Structural element of sperm axoneme in outer dense fiber[66]; UDP-N-acetylglucose-aminepyrophosphorylase activity[67]	Infertility related[12]	Testis-selective (also in muscle and liver)[12]; smooth muscle[57]; ubiquitous[51]	Upregulated in prostate and renal cancers,[51] prostate cancer, lymphoma and other cancers[53]
SPAG4	SPAG4	Role in protein localization to sperm axoneme and outer dense fibers[13,68]	—	Testis-specific[68]; testis and pancreas (little in, pituary, ovary, duodenum, lymph node)[69]; testis, and pancreatic islets[57]	Upregulated in liver, prostate, breast etc. cancer cell lines,[69] renal, lung, and other cancers[53]
SPAG5	SPAG5	Interacts with sperm axoneme and outer dense fiber structural proteins[14]; mitotic progression, centrosome integrity[70–74] and embryonic development of testis[75]	—	Testis-selective (also in thymus, pancreas, liver)[15]; testis-specific[14]; ubiquitous[51,70]	Upregulated in lung, bladder and other cancers[53]
SPAG6	SPAG6	Sperm axoneme integrity, element of flagellum central apparatus[17]; sperm motility, component of ependymal cilia,[76] ciliogenesis in bronchial epithelium cells[77]	Infertility related[17]	Testis-selective (also in lung)[17,57]; testis, oviduct, bronchial, tracheal, nasal epithelial cells[55]	Upregulated in bone marrow of AML patients,[78] prostate, colon and other cancers,[53] spinal cord neoplasm[55]
SPAG7	SPAG7	Possible structural element of sperm acrosome[18]	—	Ubiquitous[51]	Upregulated in synovial sarcoma,[79] brain, prostate and other cancers[53]
SPAG8	SPAG8	Element of sperm acrosome[19,80,81]; germ cell differentiation[82]; cell division regulation during spermatogenesis[83]	Infertility related[19,81,84]; cancer related[31]	Testis-specific[19,85]; fallopian tube, bronchial, tracheal epithelial cells[55]	Upregulated in cervical carcinoma,[32] lobular breast carcinoma, lung, and other cancers[53]
SPAG9	SPAG9	Positive regulator of JNK p38 MAPK signaling module, shuttling of preorganized signaling complexes[37,86–88]; male germ cell development, fertility, element of sperm acrosome[37,89,90]; cell migration[91]; membrane trafficking[92]; endodermal differentiation of stem cells[93]	Infertility related[20]; autoimmune disorder-related[86,94]; cancer related[33–37]	Testis-specific[20]; testis-selective alternative splice variant (also in fibroblasts, stomach, brain), total SPAG9—ubiquitous[86]	Upregulated in renal cell carcinoma[33], breast cancer,[34] epithelial ovarian cancer,[37] cervical carcinoma,[35] renal, brain and other cancers,[53] thyroid cancer[55]
SPAG10‡	MFGE8	Apoptotic cell clearance, sperm-egg binding, maintenance of epididymal and intestinal epithelia, mammary gland development, promotion of vascularization, exocytosis, facilitation of antigen presenation[95]	Cancer related[21,96,97]	Unbiquitous[98,99]	Upregulated in breast cancer,[21] breast, colon, and other cancers[53]
SPAG11	SPAG11B	Sperm binding and maturation[100–102]; antibacterial activity[103,104]	—	Testis-selective (also in epididymis, prostate, ovary)[22,104,105]	Upregulated in breast cancer, testicular seminoma, and other cancers[53]
SPAG12§	NHP2L1	Fertilization, element of sperm tail, midpiece, and postacrosome[106]; splicesome component[107]	Infertility related[108]	Testis-specific[109]; ubiquitous[107]	Upregulated in renal, bladder and other cancers[53]

TABLE 2. (*continued*)

SPAG Gene	Official Name	Evidence of Functions	Immunogencity	Expression Evidence	
				Normal Tissues*	Tumors†
SPAG13‖	SSFA2	Early cleavage of the fertilized oocyte[24]; actin regulation[110]; signal transduction[111]; energy homeostasis[112]	Infertility related[113]	Ubiquitous[110,111]	Upregulated in colon cancer,[110] brain, esophagus, and other cancers[53]
SPAG15	SPAM1	Sperm-egg adhesion[114,115]; cumulus penetration, hyaluronidase activity[116]; sperm maturation, element of sperm axoneme[117,118]; fluid reabsorption in kidney[119]; epididymosome, and uterosome component[120,121]; intracellular signaling[122]; angiogenesis[123]; increased metastatic potential of breast cancer[42]	Infertility related[124–126]	Testis-specific[25,127]; kidney, macrophages[119,128]; epididymis[129]; vagina, uterus, oviduct,[130] breast[42]; endothelium[131]; testis-selective (also in duodenum, small intestine, colon, bronchus)[51]	Upregulated in breast cancer,[43] laryngeal carcinoma,[44,45] colon cancer, melanoma, and glioblastoma cell lines,[123] lung cancer,[51] brain cancers[53]
SPAG16	SPAG16	Postmeiotic germ cell viability, sperm motility[132,133]; central apparatus element of cilia, and sperm flagella[26,134]	—	Testis-selective (also in prostate, spleen, ovary, thymus)[117]; testis-specific transcript SPAG16-L and testis-selective transcript SPAG16-S (also in trachea, brain, liver, kidney)[134]; fallopian tube, bronchial, tracheal epithelial cells, kidney[55]	Upregulated in ALL and breast and other cancers,[53] breast and urothelial cancer[55]
SPAG17	SPAG17	Sperm motility, central apparatus element of cilia and sperm flagella[135]	—	Testis-selective (also in brain, uterus, oviduct, lung)[135]; bronchial, tracheal epithelial cells[55]	Upregulated in pancreatic, brain, and other cancers[53]

*The major anatomical sites of expression reported by published literature as well as Human Protein Atlas,[51] BioGSP,[57] and Genevestigator[55] tools are named.

†Only the upregulation in cancer is summarized and, besides the literature data, the top two tumor types sorted by overexpression gene rank in the Oncomine database,[53] or by highest relative overexpression in Genevestigator[55] tool are named.

‡Owing to the vast amount of publications on SPAG10 only the review paper is cited.

§SPAG12 together with FA-1 are alternative names of the official gene symbol NHP2L1 in the current version of NCBI Gene database. The initial publications on FA-1 protein indicate its testis-specificity[108] and there is no precise cloned sequence of FA-1 as the previously published one[23] has been discontinued due to failure to prove existing as stated in the NCBI OMIM description of SSFA1—another alternative name for FA-1 (Scott, A. F. Personal Communication. Baltimore, MD, 2.8.2007). Hence it is not clear whether the initially described FA-1 protein is indeed the same as the ubiquitously expressed NHP2L1.

‖SPAG13 together with CS1 are alternative names of SSFA2 in the current version of Entrez Gene database, however, CS1 is also an alternative name of a completely distinct protein—the natural killer cell receptor SLAMF7 and should not be confused here.

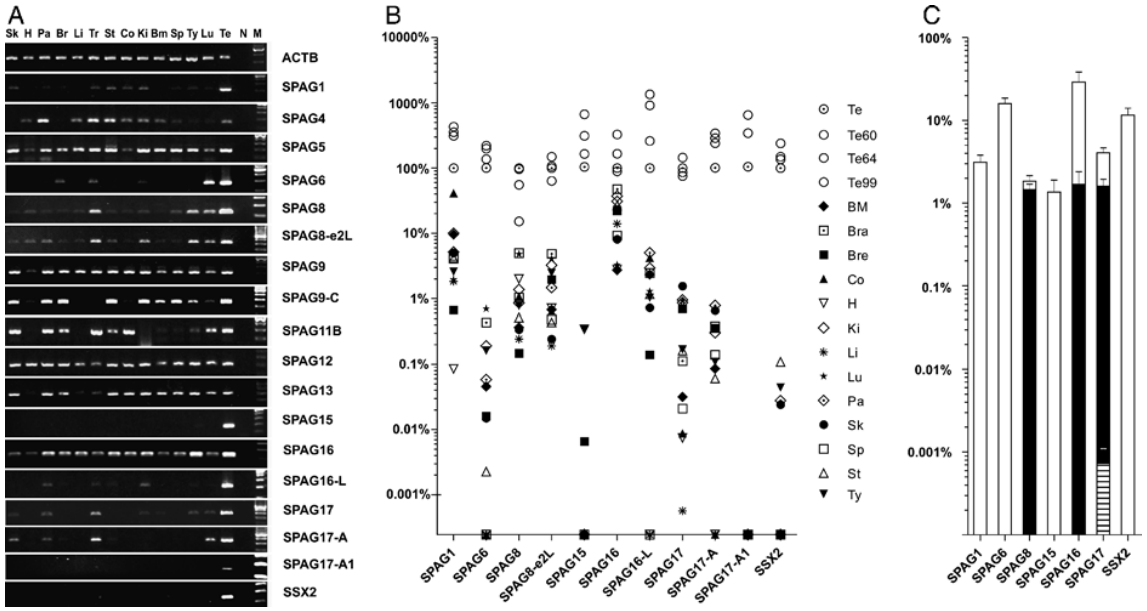


FIGURE 1. The mRNA expression of sperm-associated antigens (SPAG) in various normal tissues. A, RT-PCR was carried out on a set of 14 various normal tissues and the amplification products were visualized in a 1.2% agarose gel. B, qPCR was done on a set of 13 various normal tissues and testis samples from 4 different individuals. Mean quantities are displayed. Y axis represents percentage of the testis expression level in logarithmical scale. C, The average quantity of total SPAG (white bars) and corresponding splice variants obtained from 4 different testis samples is displayed. Black bars—SPAG8-e2L, SPAG16-L, SPAG17-A, striped bar—SPAG17-A1. Y axis represents percentage of the ACTB expression level in logarithmic scale. Error bars represent standard error of the mean. BM indicates bone marrow; Bra, brain; Bre, breast; Co, colon; H, heart; Ki, kidney; Li, liver; Lu, lung; M, size marker; N, no template control; Pa, pancreas; Sk, skin; Sp, spleen; St, stomach; Te, testis (60, 64, 99 designate 3 additional testis samples); Tr, trachea; Ty, thymus.

average of 70 empty phage controls. Statistical significance was calculated using χ^2 test. To validate SPAG17-A1 splice variant-specific antibodies, an antigen array comprising phage particles expressing SPAG17-A, SPAG17-A1, an unrelated-antigen HORMAD1, and empty phages was tested with serial 3-fold dilutions of 2 SPAG17-A1 positive gastric cancer sera. The normalized values were further normalized against StrepII signal detected by anti-Sterp II tag antibody (StrepMAB-Immo, IBA, Germany) to correct for copy number variations of recombinant proteins per phage particle.

RESULTS

Selection of Candidate CT Genes

Some of the SPAG genes have been studied extensively whereas for others only the coding sequence is known, hence, we carried out data mining in Human Protein Atlas,[50,51] Oncomine,[52,53] Genevestigator,[54,55] BioGSP,[56,57] and Entrez Gene[58] databases and the published literature to select CT gene candidates for experimental validation. The criteria for selecting a SPAG gene for further expression analyses were CT-associated expression profile, possible functional implications in oncogenesis and the cell surface localization. The decision was based on the obtained information summarized in Table 2 and included SPAG1, SPAG4, SPAG5, SPAG6, SPAG8, SPAG9, SPAG11B, SPAG15, SPAG16, and SPAG17. SPAG12 and SPAG13 were included to affirm the ubiquitous expression represented by the above mentioned online resources.

All SPAG genes are alternatively spliced as evidenced by the NCBI AceView database.[136] For the expression analysis, primers were designed to amplify a common region of all known transcripts. Additional sets of primers were designed to amplify earlier reported testis-associated splice variants of SPAG9[86] (designated as SPAG9-C), SPAG16[134] (designated as SPAG16-L), SPAG8 isoform earlier found to elicit antibody responses in melanoma patients[31] (designated as SPAG8-e2L), and SPAG17 splice variant containing the predicted antigenic region (designated as SPAG17-A). In addition, we identified novel transcripts during the cloning of the antigenic regions of SPAG6 and SPAG17 and designated them as variations of corresponding transcripts in the AceView nomenclature namely SPAG6-A1 and SPAG17-A1 (Figure, Supplemental Digital Content 2, http://links.lww.com/JIT/A85 for all alternative transcripts, amplicon positions, and cloned antigenic regions).

mRNA Expression Pattern in Normal Tissues

To describe the expression pattern of SPAG genes, we use the terms defined by Hofmann et al[137] where CT genes that are expressed mostly in testis but are detectable in low levels in a few other nongerm line tissues are designated as testis-selective, and the ones present only in germ line tissues and placenta are called testis-restricted. Expression of the selected SPAG genes together with the well-known CT antigen SSX2 as a control of testis-restricted expression (according to the CTDatabase[38]) was first analyzed by qualitative RT-PCR in a panel of 14 different normal tissues (Fig. 1A). SPAG1, SPAG6, SPAG8, SPAG8-e2L,

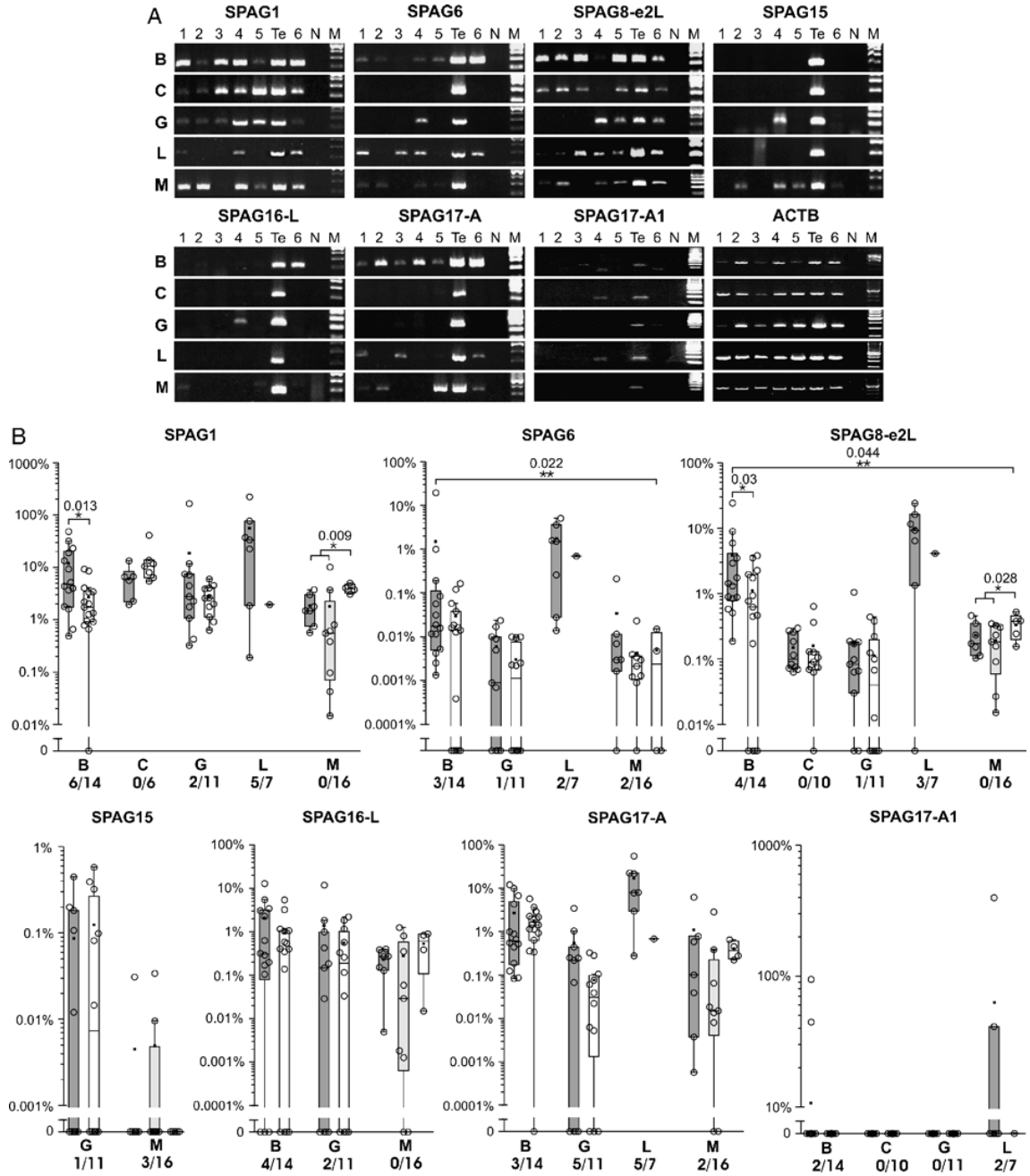


FIGURE 2. The mRNA expression of sperm-associated antigens (SPAG) in various tumors. A, RT-PCR was carried out on sets of 6 breast (B), colon (C), gastric (G), lung (L) cancer, and melanoma (M) tumor samples. Amplification products were visualized in a 1.2% agarose gel. Numbers 1-6 designate tumor tissues of cancer patients, TE-commercial normal testis, N-no template control, M-size marker. B, qPCR was carried out on sets of breast, colon, gastric, lung cancer, and melanoma samples (dark gray boxes), 9 melanoma cell lines (light gray boxes), and adjacent normal tissues and corresponding commercial normal samples (white boxes). Boxes show median with interquartile ranges for each sample set with Tukey whiskers, black dots represent the average value of the sample set, open circles designate expression values of individual tissue samples. Y axis represents percentage of the testis expression level in logarithmic scale. Single asterisk designates statistical significance reached within 1 tumor type, 2 asterisks—across all tumor types. One-tailed normal approximated *P* values, calculated using the nonparametric Mann-Whitney *U* test, are indicated if significant. Numbers below the tumor type designation indicate the number of overexpressed tumor samples out of all analyzed.

SPAG16-L, SPAG17, and SPAG17-A showed testis-selective expression, and SPAG15 and SPAG17-A1 seemed to be testis-restricted. These were further validated by qPCR in a cDNA set composed of 13 different normal tissues and testis samples from 4 different individuals. Testis-selectivity was confirmed for all analyzed SPAG transcripts (Fig. 1B), however, SPAG1 showed considerably high expression in normal colon. A minute amount of SPAG15 was observed in thymus, heart, and breast and renders it testis-selective. SPAG17-A1 was approved as testis-restricted. The remaining SPAG4, SPAG5, SPAG9, SPAG9-C, SPAG11B, SPAG12, SPAG13, and the common region of SPAG16 were present in many normal tissues in comparable levels with that of the testis sample (Fig. 1A) and this observation was confirmed by qPCR for SPAG9-C (data not shown) and SPAG16 (Fig. 1B). The objective expression level of testis-selective SPAG gene transcripts was determined in 4 different testis samples (Fig. 1C) by comparing it with the level of ACTB. It is shown that the most prominently expressed SPAG gene in testis is SPAG16 reaching about 30% of the ACTB level, whereas the testis-selective transcript SPAG16-L comprises around 6% of the total SPAG16. The transcripts containing the predicted antigenic regions of SPAG8 (SPAG8-e2L) and SPAG17 (SPAG17-A) represent approximately 80% and approximately 40% of the corresponding total mRNA, respectively, whereas the testis-restricted transcript SPAG17-A1 represents only 0.02% of the total SPAG17.

mRNA Expression in Tumor Tissues

Next, we carried out a qualitative RT-PCR screening experiment in melanoma, gastric, colon, lung, and breast cancer specimens from 6 patients in each tumor type (Fig. 2A) to determine whether any of the testis-associated SPAG transcripts are present in tumors. It is evident that the mRNA pattern varies in different cancer types among SPAGs, SPAG1, and SPAG8-e2L are the most frequent, detected in about 80% of samples. SPAG6 and SPAG17-A can be detected in about 50%, and SPAG15, SPAG16-L, and the testis-restricted transcript SPAG17-A1 are the least frequent transcripts detected in 10% to 20% of analyzed tumor samples.

To confirm upregulation of the testis-associated SPAGs in the respective tumor types, their expression was examined by quantitative RT-PCR in larger panels of tumor and adjacent normal tissue pairs. The expression level in tumor samples was compared with the adjacent normal tissue and to the average signal detected in all corresponding normal samples including the commercial normal sample to account for the interindividual expression variations and for cases when the adjacent normal tissue was not available (such as for lung cancer biopsies and melanoma cell lines). Expression of a SPAG gene was considered upregulated in a tumor sample if it exceeded that of the adjacent normal sample by at least 2-fold and the average amount in all normal samples by at least 3-fold. We could detect overexpression of all SPAG genes in tumor samples of various tumor types, however the level of SPAG15 in overexpressed tumor samples is below 1% of the testis level (Fig. 2B). Overexpression in 12.5% melanoma samples was found only for SPAG6 and SPAG17-A, whereas no SPAG gene expression was upregulated in colon cancer sample set (Fig. 2B). All SPAG genes analyzed in breast cancer showed elevated expression level with frequencies varying from 14% (SPAG17-A1) to 43% (SPAG1), and overexpression in gastric cancer was noted for all except SPAG17-A1 from 9% (SPAG6, SPAG8-e2L) to 45% (SPAG17-A) of samples (Fig. 2B). The most prominent upregulation of SPAG genes is seen in lung cancer ranging from approximately 30% (SPAG6, SPAG17-A1) to 70% (SPAG1, SPAG17-A) (Fig. 2B).

Protein Expression Analysis

To evaluate the expression of testis-selective SPAG genes at the protein level, we used immunohistochemistry on tissue microarrays comprising various normal tissues and paired breast and lung tumor-normal and unpaired gastric tumors and normal stomach tissues. The choice of analyzed proteins was limited to the availability of commercial antibodies. SPAG6 was not detected in any of the 45 normal tissues represented on TMAs (from 2 different commercial providers in 2 repeated experiments), including lung and spermatozoa (Fig. 3A). However, its expression was observed in 7 out of 12 breast cancers and 1 adjacent normal breast specimen (Fig. 3B, normal tissue panel). All positive tumor samples showed distinct perinucleolar staining (Fig. 3B, arrows in the tumor images) and mostly weak nuclear staining with strong signals in 2 out of 7 tumors (Fig. 3B). A more prominent signal was observed in 11 out of 12 lung cancer specimens strongly staining either nucleus alone or also cytoplasm and little amount in all adjacent normal lung samples with exclusively nuclear localization (Fig. 3C) just as in the adjacent normal breast sample. The positive tumors could be divided by the proportion of the positive cells, 8 out of 11 samples showing around or less than 50% and 3 out of 11 staining close to 100% of cells (Fig. 3C), however, no correlation of tumor stage or metastatic status was noted.

SPAG8 was observed only in discrete cells in the parabasal layer of ectocervix (Fig. 3D), acinar, and ductal epithelium of the breast (Fig. 3E, normal tissue panel) and stomach fundus glands (Fig. 3F), whereas all other normal tissues including testis, skin, and other stratified squamous cell epithelia such as that of esophagus or larynx were negative (data not shown). The subcellular localization is mostly cytoplasmic, but in the cells of fundus glands also the nucleus is stained and in ectocervix also possibly the plasma membrane. Its expression was also detected in 60% of breast tumors (7 out of 12) with a strong staining characteristic to nonmetastatic lesions (3 out of 3), whereas metastatic breast carcinomas (assessed by the presence of dissemination to lymph nodes) were less prominently stained (4 out of 9) and showed a possible surface localization in some cells (2 out of 4 positive cases) (Fig. 3D, arrows in tumor panel pictures). SPAG8 in gastric cancer (Fig. 3E) was only rarely detected by a strong signal and a decreased frequency of positive cases was noted with advancing stage: from 55% of stage II (11 out of 20) to 44% of stage III (8 out of 18). All inflammatory, metaplastic, and dysplastic samples and a single stage I case were positive and the single stage IV case and 10 lymph node metastases were negative. Surface localization was suspected in 1 case (gastric adenocarcinoma, stage II) (Fig. 3E, arrow in tumor image). We observed no SPAG8 protein expression neither in 40 malignant melanoma tissues nor in 8 unpaired normal skin samples (data not shown).

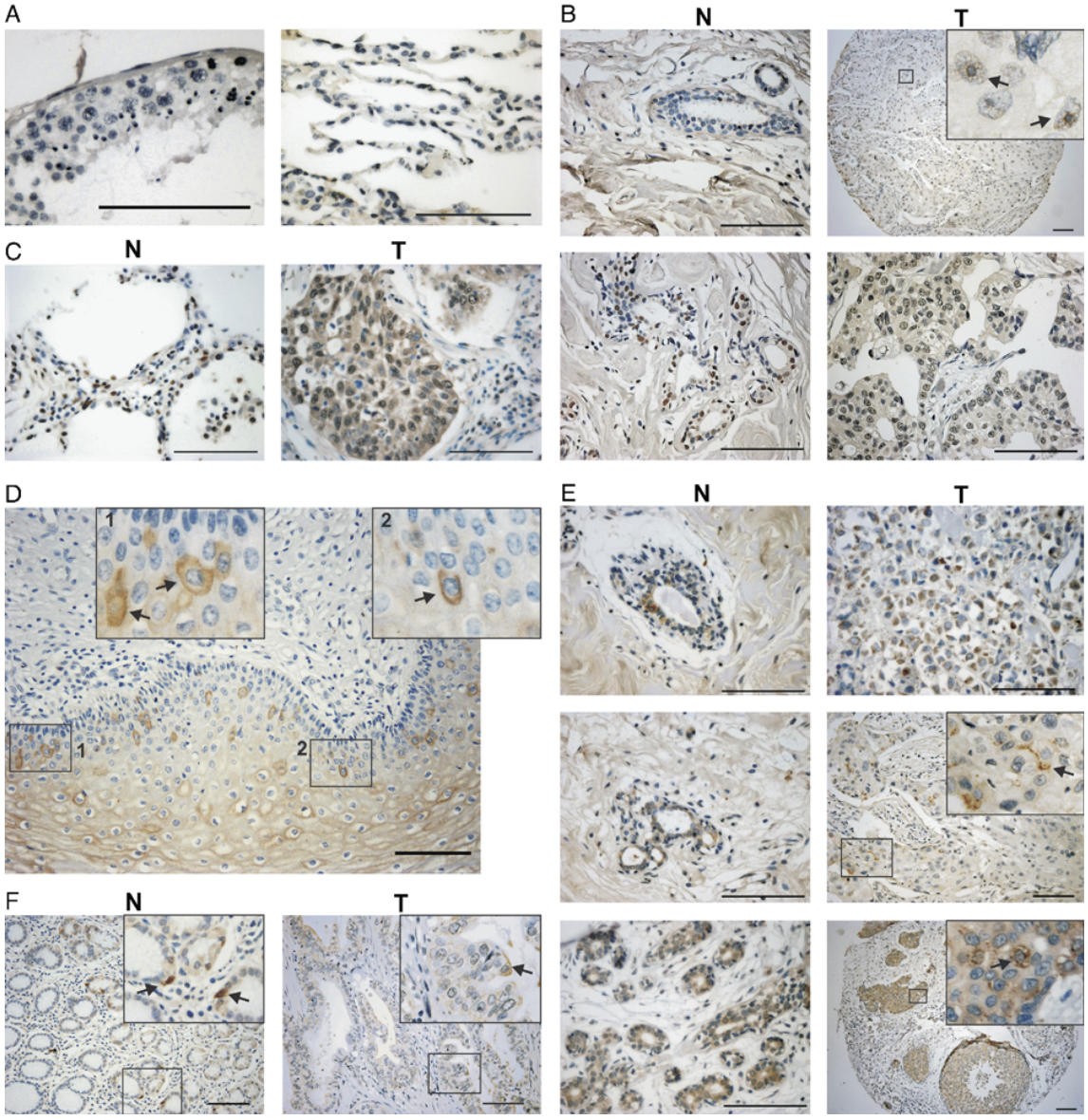


FIGURE 3. Protein expression of sperm-associated antigens (SPAG) 6 and 8. Immunohistochemical (IHC) analysis was done on paraffin-embedded tissue microarrays using anti-SPAG6 and anti-SPAG8 antibodies, HRP-conjugated secondary antibodies, DAB color development, and hematoxylin counterstaining. Black lines designate scale bars of 10 μm. A, SPAG6 IHC analysis in normal testis (left) and normal lung (right). B, SPAG6 IHC staining in normal breast (N) and paired tumor (T) tissues; arrows indicate cells with perinucleolar staining. C, SPAG6 IHC staining in normal lung (N) and paired tumor (T) tissues. D, SPAG8 IHC staining in normal stratified squamous cell epithelium of ectocervix; arrows indicate discrete strongly stained cells with possible surface staining. E, SPAG8 IHC staining in normal breast (N), and paired tumor (T) tissues, arrows indicate tumor cells with possible surface staining. F, SPAG8 staining in normal stomach (N), arrows indicate discrete strongly stained cells, and unpaired gastric tumor (T) tissues, arrows indicate tumor cells with possible surface staining.

Frequency of Autoantibodies

To analyze the immunogenicity of the cancer-associated SPAG proteins and the earlier described cancer serum marker SPAG9-C[33–37] in cancer patients, we determined the frequency of IgG class autoantibodies in sera from 539 cancer patients, 127 patients with inflammatory gastrointestinal disorders and 147 cancer-free individuals by using custom phage-displayed antigen microarray. The frequency of cancer autoantibodies against SPAG1, SPAG6-A and A1, SPAG9-C, and the novel testis-restricted splice variant SPAG17-A1 are comparable with the CT antigen HORMAD1 and higher than against MAGEA1 and SSX2, whereas no sera reacted with SPAG6-B and SPAG17-A transcripts (Table 3, Fig. 4A).

TABLE 3. The Frequency of Reactive Sera Against SPAG Proteins in Cancer Patients, Gastritis Patients, and Healthy Individuals*

Antigen	B (39)	C (33)	G (172)	L (24)	LE (28)	M (163)	P (52)	T (28)	All Ca (539)†	GI (127)	HD (147)
SPAG1	0	3	0	4.2	0	0	0	0	0.4	0	0
SPAG6-A	0	0	0	0	0	0	2	0	0.2	0	0
SPAG6-A1	0	0	0.6	0	0	0.6	0	0	0.4	0	0
SPAG6-B	0	0	0	0	0	0	0	0	0	0	0
SPAG8-e2L	0	0	1.7	0	3.6	3.1	2	3.6	2	3.1	0.7
SPAG9-A	0	0	1.2	0	0	0	0	0	0.4	0	0.7
SPAG9-C	0	0	0	0	0	0.6	0	0	0.2	0	0
SPAG16-L	0	0	0.6	0	3.6	0.6	5.8	0	1.1	3.1	2
SPAG17-A	0	0	0	0	0	0	0	0	0	0	0
SPAG17-A1	0	0	1.7	0	0	0.6	0	0	0.7	0	0
CTAG1B	10	6	7	12.5	7	10.4	9.6	10.7	8.9	0	0.7
HORMAD1	0	0	3.5	0	3.6	0.6	0	0	1.5	1.6	0
MAGEA1	0	0	0	0	0	0.6	0	0	0.2	0	0
SSX2	0	0	0	0	0	0.6	2	0	0.4	0	0

*The number of tested sera is indicated in brackets under the tested tumor type, and the frequency of reacting sera is indicated in percents.
†All Ca—frequency of antibody responses across all tumor types.
B indicates breast cancer; C, colon cancer; Ca, cancer; G, gastric cancer; GI, gastrointestinal inflammatory diseases; HD, healthy donors; L, lung cancer; LE, lymphocytic leukemia; M, melanoma; P, prostate cancer; T, thyroid cancer.

SPAG9-A and SPAG16-L were detected also by sera from healthy donors and gastritis patients in comparable frequencies with the tumor sera suggesting an inflammation rather than cancer-related response (Table 3). The most often recognized was SPAG8-e2L with high titer antibodies (antibody signal over 3) in 1 out of 167 melanoma sera (stage I) and low titer in 4 out of 167 melanoma, 3 out of 172 gastric, 1 out of 28 lymphocytic leukemia, 1 out of 52 prostate cancer, and 1 out of 28 thyroid cancer sera (stages I-III), and 4 out of 127 gastrointestinal disease patients (3 atrophic gastritis and 1 duodenum ulcer) and 1 out of 147 healthy donors (Fig. 4A, Table 3). Cancer-specific autoantibodies against SPAG17-A1 were detected in 3 out of 172 gastric cancer patients (1 stage III and 2 stage IV patients), whereas the major isoform SPAG17-A was very weakly recognized with the corresponding sera antibody signals just below the serum positivity threshold (Fig. 4A). SPAG17-A1 isoform lacks 171 amino acids compared with SPAG17-A (Table and Figure, Supplemental Digital Contents 1 and 2, http://links.lww.com/JIT/A84 and http://links.lww.com/JIT/A85) and hence, might be represented on the phage surface in a more efficient way leading to increased antibody signals. To exclude the possibility that differential serum reactivity against these 2 splice isoforms is owing to variable amount of recombinant protein on the phage surface and to determine the titer of the anti-SPAG17-A1 autoantibodies, we carried out additional screening with serial 3-fold dilutions of 2 positive gastric cancer sera on a separate array containing 5 replicates of SPAG17-A, SPAG17-A1, and irrelevant antigen HORMAD1 and the antibody signal intensities were normalized not only against the total amount of printed phage by anti-T7 phage tail antibody, but also against the copy number of recombinant protein per phage particle by anti-Strep Tag antibody (Materials and Methods). The presence of the anti-SPAG17-A1-specific antibodies was approved and the antibody signals are detected at the sera dilution higher than 1:2700 (Fig. 4B).

DISCUSSION

The specificity and oncogenicity of a potential cancer antigen are the dominant criteria for choosing it for further evaluation of immunogenicity and therapeutic functions as defined in the recent guidelines of prioritizing antigens for immunotherapy.[5] We analyzed online expression databases and published literature to define those SPAG genes that might match the above criteria and to select the candidate CT genes for experimental validation. Five of the SPAG genes (SPAG2, SPAG7, and SPAG10, SPAG12, SPAG13) were presented as ubiquitous by literature data and data deposited in online gene array databases, and this expression pattern was experimentally confirmed for SPAG12 and SPAG13. Qualitative mRNA expression analysis showed that SPAG4 is most abundant in pancreas fitting the earlier published results[69] and several other normal tissues, hence, while being a possible marker of various neoplasias,[69] it cannot be included in the CT gene category. SPAG5 was ubiquitously expressed differing from the earlier reported testis-selective distribution,[15,53] but in concordance with its functional involvement in centrosome integrity[70–74] and the protein expression pattern presented in HPA.[51] Among the most studied is SPAG9 or JLP, a scaffolding protein that is involved in various signaling events along the MAPK[87] and TNF-α or NF-κB[138] pathways and is important for sperm development[89] and retinoic acid-mediated endodermal differentiation.[93] In addition, the overexpression of a distinct testis-associated splice variant of SPAG9 (designated by AceView database[136] as SPAG9-C) is suggested to be important for tumorigenesis and proposed as a tumor immunotherapy target.[33–35,37] We show, however, that, despite the dominant expression in testis, SPAG9-C is present in various normal tissues in comparable levels and it cannot be considered as a CT antigen with implications in immunotherapy.

We continued the expression analyses of testis-selective SPAG genes in various tumor types and showed upregulation in cancer tissues of all except SPAG15. Earlier, SPAG15 has been shown as overexpressed in around 60% of breast cancer cases,[43] as well upregulation is noted in lung cancer by Oncomine,[53] however, no expression was detected in these tumors in our sample set. A larger sample size and/or other tumor types should be analyzed to see whether SPAG15 can be classified as a CT gene. According to the results of our expression analysis

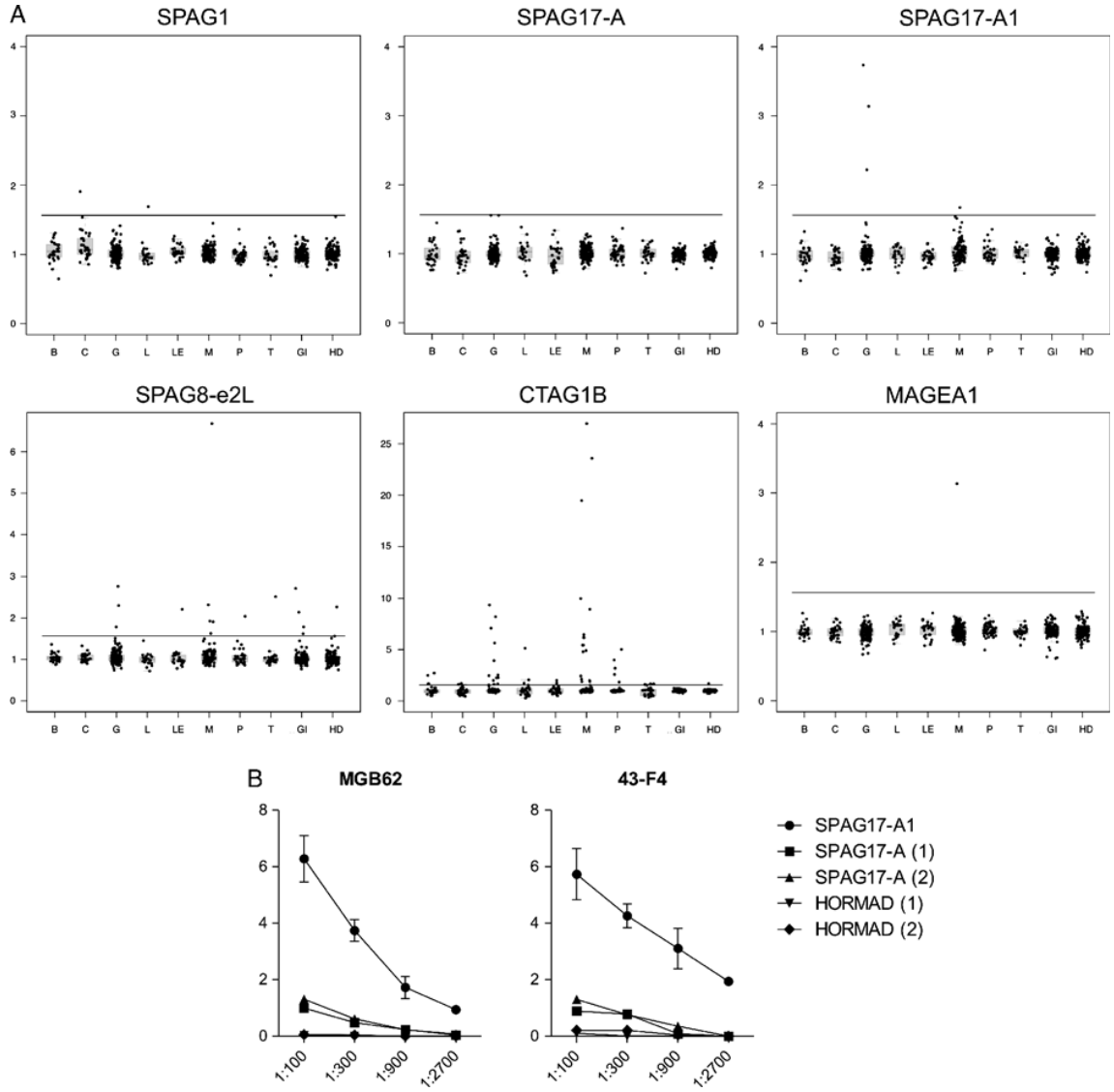


FIGURE 4. Autoantibody responses in sera of various cancer patients, gastritis patients, and healthy donors. A, Autoantibody reactivity against SPAG proteins and the well-known CT antigens was determined by screening phage-displayed antigen microarray with sera from 39 breast (B), 33 colon (C), 172 gastric (G), 24 lung cancer (L), 163 melanoma (M), 28 lymphocytic leukemia (LE), 28 thyroid (T), 52 prostate cancer (P) patients and 127 patients of gastrointestinal inflammatory diseases (GI), and 147 healthy donors (HD). Normalized average values of triplicate recombinant phages are displayed. The serum positivity threshold is defined as 4 standard deviations above the average signal of 70 nonrecombinant phage controls and is designated as a line across the graphs. B, Two gastric cancer sera MGB62 and 43-F4 showing strong reactivity against the novel splice variant SPAG17-A1 and not the dominant isoform SPAG17-A were used in serial 3-fold dilutions for screening of antigen microarray containing these phages and an irrelevant antigen HORMAD1. The obtained signals were normalized against the total amount of printed phage and the amount of recombinant surface protein per each phage. Y axis represents autoantibody signal values, X axis represents serum dilution, error bars designate standard deviations of 5 recombinant phage replicates. Average antibody signals for clone of SPAG17-A1, 2 clones of SPAG17-A, and 2 clones of irrelevant antigen HORMAD1 are displayed.

SPAG1, SPAG6, SPAG8, and SPAG17 genes are the new members of the CT gene category, whereas only distinct isoforms of SPAG16 (SPAG16-L) are testis-selective.

Tumor antigen-targeted immunotherapy trials have shown that the expansion of antigen-specific CTLs does not necessarily correlate with tumor regression,[139] whereas the proper activation of antigen-specific T helper cells has a great potential to underlie tumor control and eradication.[140] The presence of specific class switched autoantibodies in patients' sera against tumor antigens is indicative of spontaneous cell activation against the particular antigen and the detection of such autoantibodies can be used to

monitor the frequency of such responses in cancer patients. To do this, we used the phage-displayed antigen microarray system developed in our lab[31] and created a custom SPAG antigen microarray for serum screening of various cancer patients. Overall, the frequency of autoantibodies against SPAGs is low, yet higher than that of the well-known CT antigens such as MAGEA1 and SSX2. Humoral response against SPAG1, SPAG6, and the novel testis-restricted splice variant SPAG17-A1 was specific to cancer patients, and cancer-associated in the case of SPAG8 ascribing these genes as novel members of the CT antigen group, while anti-SPAG16-L antibodies are equally present in healthy individuals and gastritis patients leaving its CT antigen status in question.

SPAG1, a protein involved in G protein coupled receptor signaling during spermatogenesis and fertilization[59,60] and sperm mtDNA degradation in zygote,[62,63] has been earlier shown to be a progression marker of pancreatic cancer and a cell motility factor.[29] We, for the first time show here that SPAG1 can be immunogenic and is up-regulated prominently in lung and breast cancers, however, relatively high expression in normal colon was also detected, which correlates with gene array results presented by the Genevestigator tool,[55] noting that its application for immunotherapy might not be straightforward.

The fresh appreciation of an organelle present on most mammalian cells—the nonmotile primary cilium—as an important signal transduction hook-up has lead to the notion that centrosome and basal body (the nucleation center of ciliary microtubules) are interexchanging structures that respond to cell cycle regulation, and that cilium is necessary for the proper function of such crucial signaling pathways as the Hegdehog, Wnt and PDGFalfa, disruption of which results in various developmental disorders and cancer.[141] The proteome of primary cilium has now extended to more than 2500 proteins, many of which are shared between basal body and centrosome and are involved in cell cycle checkpoints.[142,143] It is interesting to note that 8 of the 15 SPAG genes: SPAG2, SPAG4, SPAG5, SPAG6, SPAG8, SPAG15, SPAG16, and SPAG17 have been shown to participate in centrosome and/or cilium-related events.

Initially identified as a sperm acrosome protein recognized by sperm agglutinating antibodies and necessary for sperm-egg binding,[19,80] SPAG8 has also been shown to participate in G2/M phase regulation delaying the exit form mitosis when overexpressed in CHO-K1 cells and colocalizes with microtubule organizing center (MTOC) in prophase and spindle microtubules during metaphase.[83] MTOC nucleates microtubules in both, cilium (basal body) and mitotic spindle (centrosome), and it would be of interest to determine whether the disbalanced expression of SPAG8 might participate in ciliary or mitotic defects. A relation of SPAG8 to oncogenesis comes from a study, in which cDNA microarrays showed a 5-fold overexpression in the more aggressive HPV18-type cervical carcinoma when compared with normal cervical epithelium.[32] It is interesting to note that our IHC results showed strong cytoplasmic and possibly membranal staining of distinct cells in the parabasal layer of ectocervix and, considering the frequency of the stained cells, they might correspond to the stem cells of this squamous-stratified epithelium. The basal and parabasal level of ectocervix is thought to be the location of the cervical squamous epithelium stem cells[144] and the HPV-induced cervical carcinogenesis is thought to arise from the deregulation of these stem cells.[145] It would be interesting to see if SPAG8 colocalizes to the same cells as the currently suggested cervical stem cell markers CK17 and p63[145] and whether its MTOC-related activities could be involved in the development of HPV-induced cervical cancer. In addition, we saw the staining of distinct cells in the glandular epithelia of the breast and stomach and the overexpression in the corresponding tumors was seen with a tendency to decrease with tumor stage and was not detected in any of the metastasis samples. In accordance with this observation, the autoantibodies were found mostly in sera of chronic atrophic gastritis that is an early preneoplastic condition, and sera of gastric cancer and melanoma patients with stages I to III. It is tempting to speculate that overexpression of SPAG8 could be advantageous for tumor evolvement in certain contexts but it is lost during tumor progression possibly owing to immune selection. With this in mind, autoantibodies against SPAG8 could be used as an early cancer biomarker; however, targeting of SPAG8 in tumor immunotherapy might be jeopardized. Nevertheless, the possible relation to stem cell functions and surface localization might provide another axis of treatment approaches using this antigen.

First described in *Chlamydomonas flagellum* SPAG6, SPAG16-L, and SPAG17 have been shown to mutually interact at the central apparatus of sperm axoneme—a central duplet of microtubules that is characteristic to the motile cilia, but is absent from nonmotile primary cilia, and are necessary for flagellar motility.[117,135] The expression of SPAG6 has been described in tracheal and bronchial epithelium[77] and detected in other motile cilia-covered epithelia of the respiratory tract as presented by gene array data,[55] in addition, it has been shown as a dynamically exchanging basal body component.[146] A double knockout mouse model of SPAG6 and SPAG16-L showed early mortality of litters owing to severe phenotypes of hydrocephalus and pneumonia indicating to their importance in proper functioning of mucus and fluid-propelling motile cilia, but not polycystic kidneys or left-right axis defects characteristic to nonmotile and nodal ciliopathies.[147] It is not known how these proteins contribute to the motility of mammalian cilia, as the ultrastructure of axonemes in the deficient animals is normal.[147]

Only a few studies have related these proteins to cancer. SPAG6 was reported to be overexpressed in the bone marrow of AML patients and suggested as a marker for minimal residual disease and relapse,[78] whereas SPAG17 has been suggested as a potential candidate gene of thyroid cancer susceptibility.[148] We have shown the overexpression of SPAG6 mRNA in breast and lung cancer specimens and verified this by IHC analysis. We also detected SPAG6 protein in 1 out of 12 adjacent normal breast samples and all adjacent normal lung alveoli samples, while none of these tissues from cancer-free individuals were positive. Whether the presence of SPAG6 in these adjacent normal samples represents a factor of early premalignant transformation or is a particular tumor microenvironment-induced phenomenon, remains to be determined. Taking into account the functional involvement of SPAG6 in the motile cilia of the brain and the respiratory tract, it's targeting in cancer immunotherapy raises caution. Our IHC results showed, however, that the overexpressed SPAG6 protein in breast and lung cancer had perinucleolar, nuclear, and cytoplasmic subcellular localization patterns, indicating to other possible functions

of SPAG6 besides participation in the motile cilium. The observed humoral immune response in cancer patients against transcripts of SPAG6 might also suggest that its ectopic overexpression can be immunogenic. Further studies are warranted to reveal the functional significance of SPAG6 overexpression and to define the molecular alterations underlying its ectopic expression and possibly motile cilium-unrelated activities in cancers, and might result in the identification of epitopes distinct from the normal ciliary SPAG6 providing ground for tumor-associated SPAG6 targeting.

SPAG17 showed testis-selective expression pattern and was upregulated in lung and gastric tumors, whereas its minor splice variant SPAG17-A1 was testis-restricted and present in a portion of lung and breast tumors, but it is not yet clear whether the low-level mRNA of SPAG17-A1 could result in significant quantity of protein; nevertheless, exactly this isoform is recognized by high titer autoantibodies in late-stage gastric cancer sera. The low-antibody signal against SPAG17-A in SPAG17-A1 reactive sera might represent a weak cross-reactivity between these splice variants. Testis is an organ with one of the most diverse transcriptomes owing to vastly rich alternative splicing. We and other researchers have earlier suggested that the deregulation of alternative splicing in cancer can result in the recognition of such testis-restricted splice sites, leading to production of immunogenic isoforms of otherwise tolerated proteins.[149–151] We show here for the first time a testis-restricted splice variant-specific immune response and suggest that such antigens are designated to a separate category called CT-spliced antigens. Further studies of SPAG6 and SPAG17 alternative isoform expression in various tumors and the capability to elicit protective T helper and/or CTL responses in patients bearing SPAG-positive tumors are warranted and ongoing. Furthermore, it would be of interest to find out whether these SPAG proteins might also participate in primary cilium regulated signaling pathways related to oncogenesis if their natural organelle—the motile cilium—is absent.

In conclusion, we have determined the expression pattern of SPAG genes in various normal tissues and showed that only 5 of the 15 genes in the SPAG group are actually testis-selective, although the earlier considered CT genes SPAG4 and SPAG9 are expressed in several normal tissues in comparable levels with the testis. Expression analysis in various tumor types revealed upregulation of SPAG1, SPAG6, SPAG8, the splice variants SPAG16-L, SPAG17-A, and novel testis-restricted alternative splice variant SPAG17-A1. Cancer-related humoral immune response was found against SPAG1, SPAG6, SPAG8, and SPAG17-A1, thus showing these as novel CT antigens that might be useful as cancer serum biomarkers. In addition, transcripts of SPAG6 and SPAG17-A1 are potential candidates for cancer immunotherapy.

ACKNOWLEDGMENTS

The authors thank Dr L. Ņikitina-Zaķe from the Genome Database of the Latvian population, Dr D. Schadendorf from Skin Cancer Unit in German Cancer Research Center, Dr T. Wex from the Clinic of Gastroenterology, Hepatology and Infectious Diseases, Otto-von-Guericke University Magdeburg, Dr G. Gaudernack from Norwegian Radium Hospital and Onyvax Vaccine Therapies Ltd, UK for the serum sample supply; Dr E. Liepiņš and Dr M. Dambrova from the Latvian Institute of Organic Synthesis, and Dr D. Pjanova and Dr R. Brūvere from Latvian Biomedical Research and Study Centre for excellent help in IHC and microscopy.

REFERENCES

1. Jager E, Jager D, Knuth A. Antigen-specific immunotherapy and cancer vaccines. *Int J Cancer*. 2003;106:817–820.
2. Finn OJ. Cancer immunology. *N Engl J Med*. 2008;358:2704–2715.
3. Rosenberg SA. Development of effective immunotherapy for the treatment of patients with cancer. *J Am Coll Surg*. 2004;198:685–696.
4. Reiman JM, Kmieciak M, Manjili MH, et al. Tumor immunoediting and immunosculpting pathways to cancer progression. *Semin Cancer Biol*. 2007;17:275–287.
5. Cheever MA, Allison JP, Ferris AS, et al. The prioritization of cancer antigens: a national cancer institute pilot project for the acceleration of translational research. *Clin Cancer Res*. 2009;15:5323–5337.
6. Simpson AJ, Caballero OL, Jungbluth A, et al. Cancer/testis antigens, gametogenesis and cancer. *Nat Rev Cancer*. 2005;5:615–625.
7. Jager E, Nagata Y, Gnjatic S, et al. Monitoring CD8 T-cell responses to NY-ESO-1: correlation of humoral and cellular immune responses. *Proc Natl Acad Sci U S A*. 2000;97:4760–4765.
8. Jager E, Gnjatic S, Nagata Y, et al. Induction of primary NY-ESO-1 immunity: $CD8^+$ T lymphocyte and antibody responses in peptide-vaccinated patients with NY-ESO-1+ cancers. *Proc Natl Acad Sci U S A*. 2000;97:12198–12203.
9. Jager E, Chen YT, Drijfhout JW, et al. Simultaneous humoral and cellular immune response against cancer-testis antigen NY-ESO-1: definition of human histocompatibility leukocyte antigen (HLA)-A2-binding peptide epitopes. *J Exp Med*. 1998;187:265–270.
10. Gnjatic S, Atanackovic D, Jager E, et al. Survey of naturally occurring $CD4^+$ T-cell responses against NY-ESO-1 in cancer patients: correlation with antibody responses. *Proc Natl Acad Sci U S A*. 2003;100:8862–8867.
11. Zhang ML, Wang LF, Miao SY, et al. Isolation and sequencing of the cDNA encoding the 75-kD human sperm protein related to infertility. *Chin Med J (Engl)*. 1992;105:998–1003.
12. Diekman AB, Goldberg E. Characterization of a human antigen with sera from infertile patients. *Biol Reprod*. 1994;50:1087–1093.
13. Tarnasky H, Gill D, Murthy S, et al. A novel testis-specific gene, SPAG4, whose product interacts specifically with outer dense fiber protein ODF27, maps to human chromosome 20q11.2. *Cytogenet Cell Genet*. 1998;81:65–67.
14. Shao X, Xue J, van der Hoorn FA. Testicular protein Spag5 has similarity to mitotic spindle protein deepest and binds outer dense fiber protein Odf1. *Mol Reprod Dev*. 2001;59:410–416.
15. Chang MS, Huang CJ, Chen ML, et al. Cloning and characterization of hMAP126, a new member of mitotic spindle-associated proteins. *Biochem Biophys Res Commun*. 2001;287:116–121.
16. Mack GJ, Compton DA. Analysis of mitotic microtubule-associated proteins using mass spectrometry identifies astrin, a spindle-associated protein. *Proc Natl Acad Sci U S A*. 2001;98:14434–14439.
17. Neilson LI, Schneider PA, Van Deerlin PG, et al. cDNA cloning and characterization of a human sperm antigen (SPAG6) with homology to the product of the Chlamydomonas PF16 locus. *Genomics*. 1999;60:272–280.

18. Beaton S, Cleary A, ten Have J, et al. Cloning and characterization of a fox sperm protein FSA-1. *Reprod Fertil Dev*. 1994;6:761–770.
19. Liu QY, Wang LF, Miao SY, et al. Expression and characterization of a novel human sperm membrane protein. *Biol Reprod*. 1996;54:323–330.
20. Shankar S, Mohapatra B, Suri A. Cloning of a novel human testis mRNA specifically expressed in testicular haploid germ cells, having unique palindromic sequences and encoding a leucine zipper dimerization motif. *Biochem Biophys Res Commun*. 1998;243:561–565.
21. Larocca D, Peterson JA, Urrea R, et al. A Mr 46,000 human milk fat globule protein that is highly expressed in human breast tumors contains factor VIII-like domains. *Cancer Res*. 1991;51:4994–4998.
22. Kirchhoff C, Osterhoff C, Habben I, et al. Cloning and analysis of mRNAs expressed specifically in the human epididymis. *Int J Androl*. 1990;13:155–167.
23. Naz RK, Zhu X. Molecular cloning and sequencing of cDNA encoding for human FA-1 antigen. *Mol Reprod Dev*. 2002; 63:256–268.
24. Javed AA, Naz RK. Human cleavage signal-1 protein; cDNA cloning, transcription and immunological analysis. *Gene*. 1992;112:205–211.
25. Lin Y, Kimmel LH, Myles DG, et al. Molecular cloning of the human and monkey sperm surface protein PH-20. *Proc Natl Acad Sci U S A*. 1993;90:10071–10075.
26. Smith EF, Lefebvre PA. PF20 gene product contains WD repeats and localizes to the intermicrotubule bridges in Chlamydomonas flagella. *Mol Biol Cell*. 1997;8:455–467.
27. Rupp G, O'Toole E, Porter ME. The Chlamydomonas PF6 locus encodes a large alanine/proline-rich polypeptide that is required for assembly of a central pair projection and regulates flagellar motility. *Mol Biol Cell*. 2001;12:739–751.
28. Suri A. Sperm specific proteins-potential candidate molecules for fertility control. *Reprod Biol Endocrinol*. 2004;2:10.
29. Neesse A, Gangeswaran R, Luettges J, et al. Sperm-associated antigen 1 is expressed early in pancreatic tumorigenesis and promotes motility of cancer cells. *Oncogene*. 2007;26:1533–1545.
30. Buechler S. Low expression of a few genes indicates good prognosis in estrogen receptor positive breast cancer. *BMC Cancer*. 2009;9:243.
31. Kalnina Z, Silina K, Meistere I, et al. Evaluation of T7 and lambda phage display systems for survey of autoantibody profiles in cancer patients. *J Immunol Methods*. 2008;334:37–50.
32. Vazquez-Ortiz G, Garcia JA, Ciudad CJ, et al. Differentially expressed genes between high-risk human papillomavirus types in human cervical cancer cells. *Int J Gynecol Cancer*. 2007;17:484–491.
33. Garg M, Kanojia D, Khosla A, et al. Sperm-associated antigen 9 is associated with tumor growth, migration, and invasion in renal cell carcinoma. *Cancer Res*. 2008;68: 8240–8248.
34. Kanojia D, Garg M, Gupta S, et al. Sperm-associated antigen 9, a novel biomarker for early detection of breast cancer. *Cancer Epidemiol Biomarkers Prev*. 2009;18:630–639.
35. Garg M, Kanojia D, Salhan S, et al. Sperm-associated antigen 9 is a biomarker for early cervical carcinoma. *Cancer*. 2009; 115:2671–2683.
36. Guinn BA, Bland EA, Lodi U, et al. Humoral detection of leukaemia-associated antigens in presentation acute myeloid leukaemia. *Biochem Biophys Res Commun*. 2005;335: 1293–1304.
37. Garg M, Chaurasiya D, Rana R, et al. Sperm-associated antigen 9, a novel cancer testis antigen, is a potential target for immunotherapy in epithelial ovarian cancer. *Clin Cancer Res*. 2007;13:1421–1428.
38. CTDatabase, [database online]. New York, NY, São Paulo, Petrópolis, RJ: Ludwig Institute for Cancer Research, Weill-Cornell Medical Center, Laboratório Nacional de Computação Científica; 2005. Updated 2009.
39. Ceriani RL, Sasaki M, Sussman H, et al. Circulating human mammary epithelial antigens in breast cancer. *Proc Natl Acad Sci U S A*. 1982;79:5420–5424.
40. Ceriani RL, Blank EW. Experimental therapy of human breast tumors with 131I-labeled monoclonal antibodies prepared against the human milk fat globule. *Cancer Res*. 1988;48:4664–4672.
41. Madan AK, Yu K, Dhurandhar N, et al. Association of hyaluronidase and breast adenocarcinoma invasiveness. *Oncol Rep*. 1999;6:607–609.
42. Beech DJ, Madan AK, Deng N. Expression of PH-20 in normal and neoplastic breast tissue. *J Surg Res*. 2002;103: 203–207.
43. Wang LP, Xu XM, Ning HY, et al. Expression of PH20 in primary and metastatic breast cancer and its pathological significance. *Zhonghua Bing Li Xue Za Zhi*. 2004;33: 320–323.
44. Godin DA, Fitzpatrick PC, Scandurro AB, et al. PH20: a novel tumor marker for laryngeal cancer. *Arch Otolaryngol Head Neck Surg*. 2000;126:402–404.
45. Christopoulos TA, Papageorgakopoulou N, Theocharis DA, et al. Hyaluronidase and CD44 hyaluronan receptor expression in squamous cell laryngeal carcinoma. *Biochim Biophys Acta*. 2006;1760:1039–1045.
46. Michaud EJ, Yoder BK. The primary cilium in cell signaling and cancer. *Cancer Res*. 2006;66:6463–6467.
47. Barakat MT, Scott MP. Tail wags dog: primary cilia and tumorigenesis. *Cancer Cell*. 2009;16:276–277.
48. Vandesompele J, De Preter K, Pattyn F, et al. Accurate normalization of real-time quantitative RT-PCR data by geometric averaging of multiple internal control genes. *Genome Biol*. 2002;3:research0034.1-0034.12
49. Welling GW, Weijer WJ, van der Zee R, et al. Prediction of sequential antigenic regions in proteins. *FEBS Lett*. 1985; 188:215–218.
50. Berglund L, Bjorling E, Oksvold P, et al. A genecentric Human Protein Atlas for expression profiles based on antibodies. *Mol Cell Proteomics*. 2008;7:2019–2027.
51. Human Protein Atlas, Version 6.0 [database online]. Stockholm, Uppsala, Mumbai: Royal Institute of Technology, Rudbeck Laboratory, Lab Surgpath; 2005. Updated March 26, 2010.
52. Rhodes DR, Yu J, Shanker K, et al. ONCOMINE: a cancer microarray database and integrated data-mining platform. *Neoplasia*. 2004;6:1–6.
53. Oncomine Research Edition, Version 4.3 [database online]. Ann Arbor, MI: Compendia Bioscience, Inc; 2008. Updated May 11, 2010.
54. Hruz T, Laule O, Szabo G, et al. Genevestigator v3: a reference expression database for the meta-analysis of transcriptomes. *Adv Bioinformatics*. 2008;2008:420747.
55. Genevestigator, Version 3 Zurich: Gruissem Laboratory, Widmayer Laboratory, ETH Zurich; 2008. Updated 2010.
56. Wu C, Orozco C, Boyer J, et al. BioGPS: an extensible and customizable portal for querying and organizing gene annotation resources. *Genome Biol*. 2009;10:R130.
57. BioGSP, [database online]. San Diego, CA: Genomics Institute of the Novartis Research Foundation; 2009. Updated 2010.
58. Entrez Gene, [database online]. Bethesda, MD: National Center for Biotechnology Information; 1988. Updated 2010.
59. Lin W, Zhou X, Zhang M, et al. Expression and function of the HSD-3.8 gene encoding a testis-specific protein. *Mol Hum Reprod*. 2001;7:811–818.
60. Lin W, Miao SY, Zhang L, et al. Study on the function of HSD-3.8 gene encoding a testis-specific protein with yeast two-hybrid system. *Zhongguo Yi Xue Ke Xue Yuan Xue Bao*. 2002;24:582–587.
61. Liu N, Qiao Y, Cai C, et al. A sperm component, HSD-3.8 (SPAG1), interacts with G-protein beta 1 subunit and activates extracellular signal-regulated kinases (ERK). *Front Biosci*. 2006;11:1679–1689.

62. Hayashida K, Omagari K, Masuda J, et al. The sperm mitochondria-specific translocator has a key role in maternal mitochondrial inheritance. *Cell Biol Int*. 2005;29:472–481.
63. Hayashida K, Omagari K, Masuda J, et al. An integrase of endogenous retrovirus is involved in maternal mitochondrial DNA inheritance of the mouse. *Biochem Biophys Res Commun*. 2008;366:206–211.
64. Takaishi M, Huh N. A tetratricopeptide repeat-containing protein gene, tpis, whose expression is induced with differentiation of spermatogenic cells. *Biochem Biophys Res Commun*. 1999;264:81–85.
65. Biermann K, Heukamp LC, Steger K, et al. Genome-wide expression profiling reveals new insights into pathogenesis and progression of testicular germ cell tumors. *Cancer Genomics Proteomics*. 2007;4:359–367.
66. Diekman AB, Olson G, Goldberg E. Expression of the human antigen SPAG2 in the testis and localization to the outer dense fibers in spermatozoa. *Mol Reprod Dev*. 1998;50: 284–293.
67. Mio T, Yabe T, Arisawa M, et al. The eukaryotic UDP-N-acetylglucosamine pyrophosphorylases. Gene cloning, protein expression, and catalytic mechanism. *J Biol Chem*. 1998;273: 14392–14397.
68. Shao X, Tarnasky HA, Lee JP, et al. Spag4, a novel sperm protein, binds outer dense-fiber protein Odf1 and localizes to microtubules of manchette and axoneme. *Dev Biol*. 1999; 211:109–123.
69. Kennedy C, Sebire K, de Kretser DM, et al. Human sperm associated antigen 4 (SPAG4) is a potential cancer marker. *Cell Tissue Res*. 2004;315:279–283.
70. Gruber J, Harborth J, Schnabel J, et al. The mitotic-spindle-associated protein astrin is essential for progression through mitosis. *J Cell Sci*. 2002;115(Pt 21):4053–4059.
71. Thein KH, Kleylein-Sohn J, Nigg EA, et al. Astrin is required for the maintenance of sister chromatid cohesion and centrosome integrity. *J Cell Biol*. 2007;178:345–354.
72. Liu L, Akhter S, Bae JB, et al. SNM1B/Apollo interacts with astrin and is required for the prophase cell cycle checkpoint. *Cell Cycle*. 2009;8:628–638.
73. Du J, Jablonski S, Yen TJ, et al. Astrin regulates Aurora-A localization. *Biochem Biophys Res Commun*. 2008;370: 213–219.
74. Cheng TS, Hsiao YL, Lin CC, et al. hNinein is required for targeting spindle-associated protein Astrin to the centrosome during the S and G2 phases. *Exp Cell Res*. 2007;313: 1710–1721.
75. Suzuki H, Yagi M, Suzuki K. Duplicated insertion mutation in the microtubule-associated protein Spag5 (astrin/MAP126) and defective proliferation of immature Sertoli cells in rat hypogonadic (hgn/hgn) testes. *Reproduction*. 2006;132:79–93.
76. Sapiro R, Kostetskii I, Olds-Clarke P, et al. Male infertility, impaired sperm motility, and hydrocephalus in mice deficient in sperm-associated antigen 6. *Mol Cell Biol*. 2002;22: 6298–6305.
77. Lonergan KM, Chari R, Deleeuw RJ, et al. Identification of novel lung genes in bronchial epithelium by serial analysis of gene expression. *Am J Respir Cell Mol Biol*. 2006;35:651–661.
78. Steinbach D, Schramm A, Eggert A, et al. Identification of a set of seven genes for the monitoring of minimal residual disease in pediatric acute myeloid leukemia. *Clin Cancer Res*. 2006;12:2434–2441.
79. Fernebro J, Francis P, Eden P, et al. Gene expression profiles relate to SS18/SSX fusion type in synovial sarcoma. *Int J Cancer*. 2006;118:1165–1172.
80. Cheng GY, Shi JL, Wang M, et al. Inhibition of mouse acrosome reaction and sperm-zona pellucida binding by anti-human sperm membrane protein 1 antibody. *Asian J Androl*. 2007;9:23–29.
81. Kuang Y, Yan YC, Gao AW, et al. Immune responses in rats following oral immunization with attenuated Salmonella typhimurium expressing human sperm antigen. *Arch Androl*. 2000;45:169–180.
82. Tang X, Zhang J, Cai Y, et al. Sperm membrane protein (hSMP-1) and RanBPM complex in the microtubule-organizing centre. *J Mol Med*. 2004;82:383–388.
83. Li R, Tang XL, Miao SY, et al. Regulation of the G2/M phase of the cell cycle by sperm associated antigen 8 (SPAG8) protein. *Cell Biochem Funct*. 2009;27:264–268.
84. Wu YW, Chen DH, Miao SY, et al. Eliciting an immune response by plasmid DNA encoding a human sperm protein (HSD-1). *Arch Androl*. 1999;42:127–136.
85. Zhang XD, Miao SY, Wang LF, et al. Human sperm membrane protein (hSMP-1): a developmental testis-specific component during germ cell differentiation. *Arch Androl*. 2000;45:239–246.
86. Yasuoka H, Ihn H, Medsger TA Jr, et al. A novel protein highly expressed in testis is overexpressed in systemic sclerosis fibroblasts and targeted by autoantibodies. *J Immunol*. 2003; 171:6883–6890.
87. Lee CM, Onesime D, Reddy CD, et al. JLP: a scaffolding protein that tethers JNK/p38 MAPK signaling modules and transcription factors. *Proc Natl Acad Sci U S A*. 2002;99: 14189–14194.
88. Nguyen Q, Lee CM, Le A, et al. JLP associates with kinesin light chain 1 through a novel leucine zipper-like domain. *J Biol Chem*. 2005;280:30185–30191.
89. Iwanaga A, Wang G, Gantulga D, et al. Ablation of the scaffold protein JLP causes reduced fertility in male mice. *Transgenic Res*. 2008;17:1045–1058.
90. Jagadish N, Rana R, Selvi R, et al. Characterization of a novel human sperm-associated antigen 9 (SPAG9) having structural homology with c-Jun N-terminal kinase-interacting protein. *Biochem J*. 2005;389(Pt 1):73–82.
91. Gantulga D, Tuvshintugs B, Endo Y, et al. The scaffold protein c-Jun NH2-terminal kinase-associated leucine zipper protein regulates cell migration through interaction with the G protein G(alpha 13). *J Biochem*. 2008;144: 693–700.
92. Ikonomov OC, Fligger J, Sbrissa D, et al. Kinesin adapter JLP links PIKfyve to microtubule-based endosome-to-trans-Golgi network traffic of furin. *J Biol Chem*. 2009;284: 3750–3761.
93. Kashef K, Xu H, Reddy EP, et al. Endodermal differentiation of murine embryonic carcinoma cells by retinoic acid requires JLP, a JNK-scaffolding protein. *J Cell Biochem*. 2006;98: 715–722.
94. Yasuoka H, Kuwana M. Autoantibody response against a novel testicular antigen protein highly expressed in testis (PHET) in SSc patients. *Autoimmun Rev*. 2007;6:228–231.
95. Raymond A, Ensslin MA, Shur BD. SED1/MFG-E8: a bi-motif protein that orchestrates diverse cellular interactions. *J Cell Biochem*. 2009;106:957–966.
96. Liu Y, Chiriva-Internati M, You C, et al. Use and specificity of breast cancer antigen/milk protein BA46 for generating anti-self-cytotoxic T lymphocytes by recombinant adeno-associated virus-based gene loading of dendritic cells. *Cancer Gene Ther*. 2005;12:304–312.
97. Zeelenberg IS, Ostrowski M, Krumeich S, et al. Targeting tumor antigens to secreted membrane vesicles in vivo induces efficient antitumor immune responses. *Cancer Res*. 2008;68: 1228–1235.
98. Ensslin M, Vogel T, Calvete JJ, et al. Molecular cloning and characterization of P47, a novel boar sperm-associated zona pellucida-binding protein homologous to a family of mammalian secretory proteins. *Biol Reprod*. 1998;58:1057–1064.
99. Oshima K, Aoki N, Negi M, et al. Lactation-dependent expression of an mRNA splice variant with an exon for a multiply O-glycosylated domain of mouse milk fat globule glycoprotein MFG-E8. *Biochem Biophys Res Commun*. 1999; 254:522–528.
100. Hamil KG, Sivashanmugam P, Richardson RT, et al. HE2beta and HE2gamma, new members of an epididymis-specific family of androgen-regulated proteins in the human. *Endocrinology*. 2000;141:1245–1253.

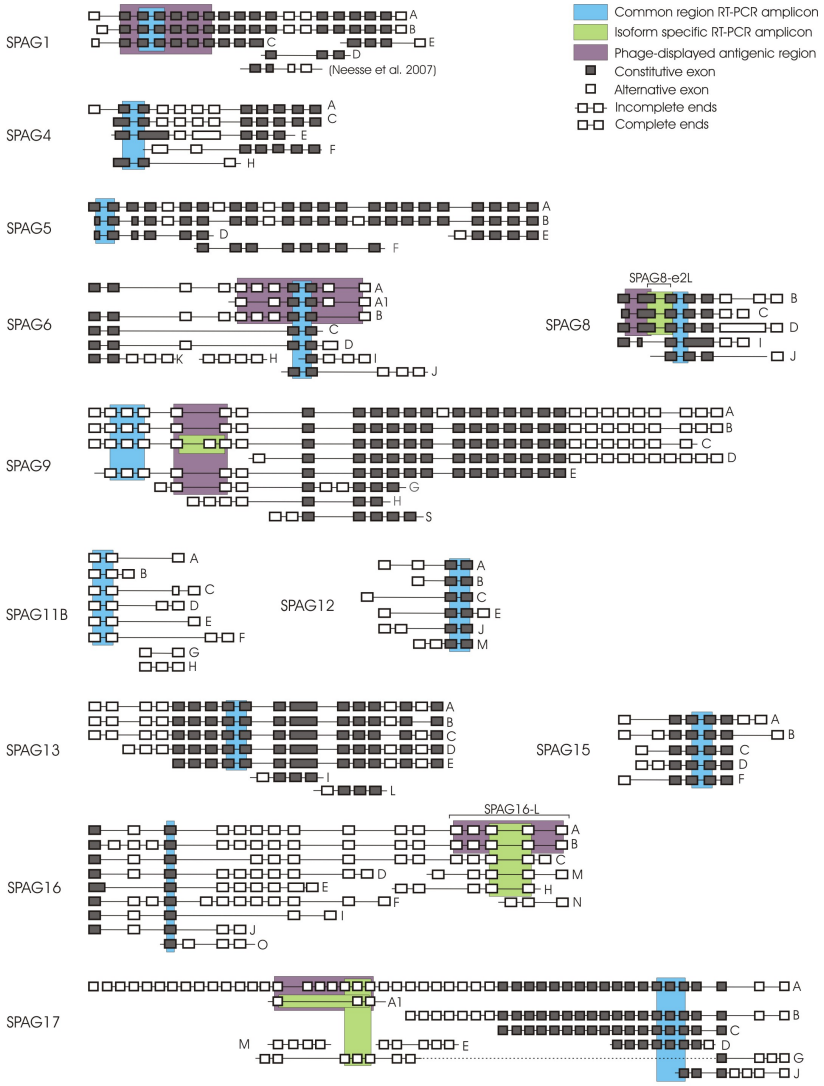
101. Osterhoff C, Kirchhoff C, Krull N, et al. Molecular cloning and characterization of a novel human sperm antigen (HE2) specifically expressed in the proximal epididymis. *Biol Reprod*. 1994;50:516–525.
102. Frohlich O, Po C, Young LG. Organization of the human gene encoding the epididymis-specific EP2 protein variants and its relationship to defensin genes. *Biol Reprod*. 2001; 64:1072–1079.
103. Yenugu S, Hamil KG, Birse CE, et al. Antibacterial properties of the sperm-binding proteins and peptides of human epididymis 2 (HE2) family; salt sensitivity, structural dependence and their interaction with outer and cytoplasmic membranes of *Escherichia coli*. *Biochem J*. 2003;372(Pt 2): 473–483.
104. Yenugu S, Hamil KG, Grossman G, et al. Identification, cloning and functional characterization of novel sperm associated antigen 11 (SPAG11) isoforms in the rat. *Reprod Biol Endocrinol*. 2006;4:23.
105. Avellar MC, Honda L, Hamil KG, et al. Novel aspects of the sperm-associated antigen 11 (SPAG11) gene organization and expression in cattle (Bos taurus). *Biol Reprod*. 2007;76: 1103–1116.
106. Naz RK. Effects of sperm-reactive antibodies present in human infertile sera on fertility of female rabbits. *J Reprod Immunol*. 1990;18:161–177.
107. Nottrott S, Hartmuth K, Fabrizio P, et al. Functional interaction of a novel 15.5 kD (U4/U6.U5) tri-snRNP protein with the 5′ stem-loop of U4 snRNA. *EMBO J*. 1999;18: 6119–6133.
108. Naz RK, Alexander NJ, Isahakia M, et al. Monoclonal antibody to a human germ cell membrane glycoprotein that inhibits fertilization. *Science*. 1984;225:342–344.
109. Zhu X, Naz RK. Fertilization antigen-1: cDNA cloning, testis-specific expression, and immunocontraceptive effects. *Proc Natl Acad Sci U S A*. 1997;94:4704–4709.
110. Inokuchi J, Komiya M, Baba I, et al. Deregulated expression of KRAP, a novel gene encoding actin-interacting protein, in human colon cancer cells. *J Hum Genet*. 2004; 49:46–52.
111. Fujimoto T, Koyanagi M, Baba I, et al. Analysis of KRAP expression and localization, and genes regulated by KRAP in a human colon cancer cell line. *J Hum Genet*. 2007;52: 978–984.
112. Fujimoto T, Miyasaka K, Koyanagi M, et al. Altered energy homeostasis and resistance to diet-induced obesity in KRAP-deficient mice. *PLoS One*. 2009;4:e4240.
113. Naz RK. Effects of antisperm antibodies on early cleavage of fertilized ova. *Biol Reprod*. 1992;46:130–139.
114. Myles DG, Primakoff P. Localized surface antigens of guinea pig sperm migrate to new regions prior to fertilization. *J Cell Biol*. 1984;99:1634–1641.
115. Griffiths GS, Miller KA, Galileo DS, et al. Murine SPAM1 is secreted by the estrous uterus and oviduct in a form that can bind to sperm during capacitation: acquisition enhances hyaluronic acid-binding ability and cumulus dispersal efficiency. *Reproduction*. 2008;135:293–301.
116. Gmachl M, Sagan S, Ketter S, et al. The human sperm protein PH-20 has hyaluronidase activity. *FEBS Lett*. 1993; 336:545–548.
117. Zhang Z, Sapiro R, Kapfhamer D, et al. A sperm-associated WD repeat protein orthologous to Chlamydomonas PF20 associates with Spag6, the mammalian orthologue of Chlamydomonas PF16. *Mol Cell Biol*. 2002;22: 7993–8004.
118. Zhang H, Martin-Deleon PA. Mouse epididymal Spam1 (pH-20) is released in the luminal fluid with its lipid anchor. *J Androl*. 2003;24:51–58.
119. Grigorieva A, Griffiths GS, Zhang H, et al. Expression of SPAM1 (PH-20) in the murine kidney is not accompanied by hyaluronidase activity: evidence for potential roles in fluid and water reabsorption. *Kidney Blood Press Res*. 2007;30: 145–155.
120. Sullivan R, Frenette G, Girouard J. Epididymosomes are involved in the acquisition of new sperm proteins during epididymal transit. *Asian J Androl*. 2007;9:483–491.
121. Griffiths GS, Galileo DS, Reese K, et al. Investigating the role of murine epididymosomes and uterosomes in GPI-linked protein transfer to sperm using SPAM1 as a model. *Mol Reprod Dev*. 2008;75:1627–1636.
122. Cherr GN, Yudin AI, Overstreet JW. The dual functions of GPI-anchored PH-20: hyaluronidase and intracellular signaling. *Matrix Biol*. 2001;20:515–525.
123. Liu D, Pearlman E, Diaconu E, et al. Expression of hyaluronidase by tumor cells induces angiogenesis in vivo. *Proc Natl Acad Sci U S A*. 1996;93:7832–7837.
124. Primakoff P, Lathrop W, Woolman L, et al. Fully effective contraception in male and female guinea pigs immunized with the sperm protein PH-20. *Nature*. 1988;335:543–546.
125. Tung KS, Primakoff P, Woolman-Gamer L, et al. Mechanism of infertility in male guinea pigs immunized with sperm PH-20. *Biol Reprod*. 1997;56:1133–1141.
126. Deng X, Meyers SA, Tollner TL, et al. Immunological response of female macaques to the PH-20 sperm protein following injection of recombinant proteins or synthesized peptides. *J Reprod Immunol*. 2002;54:93–115.
127. Jones MH, Davey PM, Aplin H, et al. Expression analysis, genomic structure, and mapping to 7q31 of the human sperm adhesion molecule gene SPAM1. *Genomics*. 1995;29:796–800.
128. Sun L, Feusi E, Sibalic A, et al. Expression profile of hyaluronidase mRNA transcripts in the kidney and in renal cells. *Kidney Blood Press Res*. 1998;21:413–418.
129. Deng X, He Y, Martin-Deleon PA. Mouse Spam1 (PH-20): evidence for its expression in the epididymis and for a new category of spermatogenic-expressed genes. *J Androl*. 2000; 21:822–832.
130. Zhang H, Martin-DeLeon PA. Mouse Spam1 (PH-20) is a multifunctional protein: evidence for its expression in the female reproductive tract. *Biol Reprod*. 2003;69:446–454.
131. Mohamadzadeh M, DeGrendele H, Arizpe H, et al. Proinflammatory stimuli regulate endothelial hyaluronan expression and CD44/HA-dependent primary adhesion. *J Clin Invest*. 1998;101:97–108.
132. Zhang Z, Kostetskii I, Tang W, et al. Deficiency of SPAG16L causes male infertility associated with impaired sperm motility. *Biol Reprod*. 2006;74:751–759.
133. Zhang Z, Kostetskii I, Moss SB, et al. Haploinsufficiency for the murine orthologue of Chlamydomonas PF20 disrupts spermatogenesis. *Proc Natl Acad Sci U S A*. 2004;101: 12946–12951.
134. Pennarun G, Bridoux AM, Escudier E, et al. Isolation and expression of the human hPF20 gene orthologous to Chlamydomonas PF20: evaluation as a candidate for axonemal defects of respiratory cilia and sperm flagella. *Am J Respir Cell Mol Biol*. 2002;26:362–370.
135. Zhang Z, Jones BH, Tang W, et al. Dissecting the axoneme interactome: the mammalian orthologue of Chlamydomonas PF6 interacts with sperm-associated antigen 6, the mammalian orthologue of Chlamydomonas PF16. *Mol Cell Proteomics*. 2005;4:914–923.
136. AceView, [database online]. Bethesda, MD: National Center for Biotechnology Information; 2000. Updated December 12, 2009.
137. Hofmann O, Caballero OL, Stevenson BJ, et al. Genome-wide analysis of cancer/testis gene expression. *Proc Natl Acad Sci U S A*. 2008;105:20422–20427.
138. Bouwmeester T, Bauch A, Ruffner H, et al. A physical and functional map of the human TNF-alpha/NF-kappa B signal transduction pathway. *Nat Cell Biol*. 2004;6:97–105.
139. Rosenberg SA, Sherry RM, Morton KE, et al. Tumor progression can occur despite the induction of very high levels of self/tumor antigen-specific $CD8^+$ T cells in patients with melanoma. *J Immunol*. 2005;175:6169–6176.
140. Muranski P, Restifo NP. Adoptive immunotherapy of cancer using CD4(+) T cells. *Curr Opin Immunol*. 2009;21:200–208.

141. Veland IR, Awan A, Pedersen LB, et al. Primary cilia and signaling pathways in mammalian development, health and disease. *Nephron Physiol*. 2009;111:p39–p53.
142. Plotnikova OV, Golemis EA, Pugacheva EN. Cell cycle-dependent ciliogenesis and cancer. *Cancer Res*. 2008;68: 2058–2061.
143. Ciliary proteome database, Version 3 [database online]. Baltimore, MD: Johns Hopkins University, McKusick-Nathans Institute of Genetic Medicine; 2006. Updated 2008.
144. Zwillenberg LO. At 40 years of the "Golden Chain". Which are the stem cells in ectocervical epithelium? *Gynecol Obstet Invest*. 1998;46:247–251.
145. Martens JE, Arends J, Van der Linden PJ, et al. Cytokeratin 17 and p63 are markers of the HPV target cell, the cervical stem cell. *Anticancer Res*. 2004;24:771–775.
146. Pearson CG, Giddings TH Jr, Winey M. Basal body components exhibit differential protein dynamics during nascent basal body assembly. *Mol Biol Cell*. 2009;20:904–914.
147. Zhang Z, Tang W, Zhou R, et al. Accelerated mortality from hydrocephalus and pneumonia in mice with a combined deficiency of SPAG6 and SPAG16L reveals a functional interrelationship between the two central apparatus proteins. *Cell Motil Cytoskeleton*. 2007;64:360–376.
148. Baida A, Akdi M, Gonzalez-Flores E, et al. Strong association of chromosome 1p12 loci with thyroid cancer susceptibility. *Cancer Epidemiol Biomarkers Prev*. 2008;17:1499–1504.
149. Kalnina Z, Zayakin P, Silina K, et al. Alterations of pre-mRNA splicing in cancer. *Genes Chromosomes Cancer*. 2005; 42:342–357.
150. Kalnina Z, Silina K, Line A. Autoantibody profiles as biomarkers for response to therapy and early detection of cancer. *Curr Cancer Ther Rev*. 2008;4:149–156.
151. Yang F, Chen IH, Xiong Z, et al. Model of stimulation-responsive splicing and strategies in identification of immunogenic isoforms of tumor antigens and autoantigens. *Clin Immunol*. 2006;121:121–133.

Supplemental Digital Content 1. Table – recombinant T7 phage displayed antigens used for microarray production.

Antigen	Length of cloned antigen, aa	Aa position in full length sequence	Genbank accession number
CTAG1B	105	1-105	NM_139250.1
HORMAD1	289	34-312	NM_032132.3
MAGEA1	129	1-129	NM_004988.4
SPAG1	215	1-215	NM003114
SPAG6-A1	235	217-451	HM370395
SPAG6-A	293	217-509	NM012443
SPAG6-B	242	217-458	NM172242
SPAG8-e2L	91	9-99	NM001039592
SPAG9-A	127	234-360	NM001130528
SPAG9-C	113	234-346	NM003971
SPAG16-L	223	409-631	NM024532
SPAG17-A	286	708-993	NM206996
SPAG17-A1	115	708-822	HM370394

Supplemental Digital Content 2. TIFF – Non-redundant alternative transcripts and RT-PCR amplification and phage-displayed antigenic regions.



The schematic structure of the unique alternative transcripts of SPAG genes deposited in the NCBI AceView database as well as transcripts identified by us (SPAG6-A1, SPAG17-A1) and other authors (SPAG1 with reference in parenthesis) is displayed. The size ratios of the exon and intron items are not proportional to their actual size ratios. An exon was considered as alternative (white boxes) if there was at least one complete transcript lacking it. The transcript nomenclature and complete or incomplete ends (approved transcription initiation and/or termination sites) are as in the AceView database. The amplification of isoform-specific regions (green boxes) was ensured by using primers in the unique exons or exon-exon border spanning primers.

3.4 Alterations of pre-mRNA splicing in cancer

GENES, CHROMOSOMES & CANCER 42:342–357 (2005)

REVIEW ARTICLE

Alterations of Pre-mRNA Splicing in Cancer

Zane Kalniņa, Pawel Zayakin, Karīna Siliņa, and Aija Linē*

Biomedical Research and Study Centre, University of Latvia, Riga, Latvia

Recent genomewide analyses of alternative splicing (AS) indicate that up to 70% of human genes may have alternative splice forms, suggesting that AS together with various posttranslational modifications plays a major role in the production of proteome complexity. Splice-site selection under normal physiological conditions is regulated in the developmental stage in a tissue type-specific manner by changing the concentrations and the activity of splicing regulatory proteins. Whereas spliceosomal errors resulting in the production of aberrant transcripts rarely occur in normal cells, they seem to be an intrinsic property of cancer cells. Changes in splice-site selection have been observed in various types of cancer and may affect genes implicated in tumor progression (for example, *CD44, MDM2,* and *FHIT*) and in susceptibility to cancer (for example, *BRCA1* and *APC*). Splicing defects can arise from inherited or somatic mutations in *cis*-acting regulatory elements (splice donor, acceptor and branch sites, and exonic and intronic splicing enhancers and silencers) or variations in the composition, concentration, localization, and activity of regulatory proteins. This may lead to altered efficiency of splice-site recognition, resulting in overexpression or down-regulation of certain splice variants, a switch in splice-site usage, or failure to recognize splice sites correctly, resulting in cancer-specific splice forms. At least in some cases, changes in splicing have been shown to play a functionally significant role in tumorigenesis, either by inactivating tumor suppressors or by gain of function of proteins promoting tumor development. Moreover, cancer-specific splicing events may generate novel epitopes that can be recognized by the host's immune system as cancer specific and may serve as targets for immunotherapy. Thus, the identification of cancer-specific splice forms provides a novel source for the discovery of diagnostic or prognostic biomarkers and tumor antigens suitable as targets for therapeutic intervention.

CONSTITUTIVE AND ALTERNATIVE SPLICING

Basal Splicing Machinery

The coding regions (exons) of most human genes are interrupted by noncoding intervening sequences (introns) that are removed from pre-mRNA molecules to produce mature mRNAs through the process of RNA splicing. RNA splicing is carried out cotranscriptionally by a spliceosome—a multicomponent complex consisting of five small nuclear ribonucleoproteins (snRNPs—U1, U2, U4, U5, and U6) and more than 100 proteins (Faustino et al., 2003). Via multiple RNA–RNA, RNA–protein, and protein–protein interactions, the spliceosome recognizes exon–intron boundaries and catalyzes two sequential trans-esterification reactions that remove introns and ligate exons (Patel et al., 2003). Intron removal must be performed with the precision of up to one nucleotide. However, it is not fully understood how it is accomplished because exon–intron borders are defined by short and weakly conserved classical splice-site sequences (Fig. 1). The 5′ and the 3′ termini of an intron are defined by a GU and an AG dinucleotide, respectively, in 95%–99% of cases (Burset et al., 2000; Black, 2003). In the 5′ splice site, the GU dinucleotide is followed by a less conserved consensus sequence, whereas the 3′ splice site consists of a branch point most often containing an adenine residue, followed by the pyrimidine-rich region and the conserved AG dinucleotide (Black, 2003). In addition to these sequences, exonic and intronic *cis* elements, known as splicing enhancers and silencers, are involved in both constitutive and alternative splicing, in which they aid in the correct identification of exon–intron borders and prevent pseudoexons from being included in the mRNA. They appear to be particularly important in the regulation of splice-site usage during alternative splicing (Cartegni et al., 2002).

Alternative Splicing

Alternative splicing (AS) of a pre-mRNA is a fundamental mechanism of differential gene expression that allows the production of structurally and functionally distinct proteins from a single coding

Supported by: Latvian Council of Science and the Royal Society.

*Correspondence to: Aija Linē, Biomedical Research and Study Centre, University of Latvia, Ratsupites St 1, LV-1067, Riga, Latvia. E-mail: aija@biomed.lu.lv

Received 29 July 2004; Accepted 4 November 2004

DOI 10.1002/gcc.20156

Published online 11 January 2005 in Wiley InterScience (www.interscience.wiley.com).

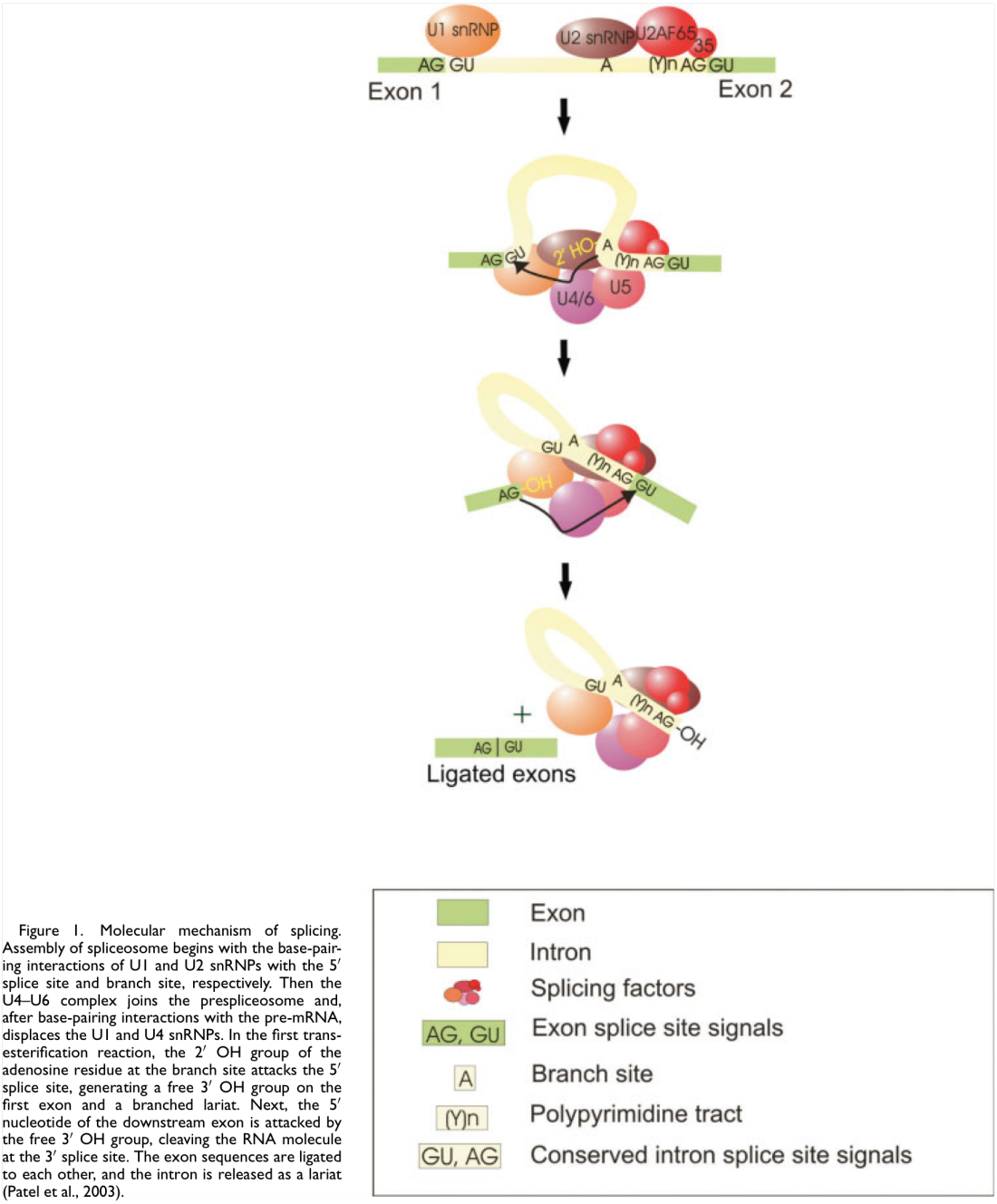


Figure 1. Molecular mechanism of splicing. Assembly of spliceosome begins with the base-pairing interactions of U1 and U2 snRNPs with the 5′ splice site and branch site, respectively. Then the U4–U6 complex joins the prespliceosome and, after base-pairing interactions with the pre-mRNA, displaces the U1 and U4 snRNPs. In the first transesterification reaction, the 2′ OH group of the adenosine residue at the branch site attacks the 5′ splice site, generating a free 3′ OH group on the first exon and a branched lariat. Next, the 5′ nucleotide of the downstream exon is attacked by the free 3′ OH group, cleaving the RNA molecule at the 3′ splice site. The exon sequences are ligated to each other, and the intron is released as a lariat (Patel et al., 2003).

sequence. The finding that the human genome contains only about 30,000–40,000 genes—only about twice as many as in the fly (Lander et al., 2001)—came as a surprise to many. Nevertheless, the proteome might number more than several hundred thousand protein species. Numerous genomewide analyses of AS have indicated that 35%–74% of human genes may have alternative splice forms (Modrek et al., 2001, 2002; Johnson et al., 2003), suggesting that AS together with various posttranslational modifications plays a major role in the production of proteome complexity.

At least five distinct alternative splicing patterns have been observed (Fig. 2): (1) the cassette exon

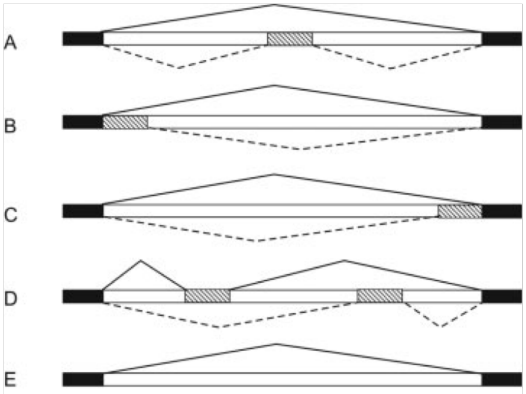


Figure 2. Patterns of alternative splicing: (A) exon skipping/inclusion; (B) alternative 5′ splice site; (C) alternative 3′ splice site; (D) mutually exclusive exons; (E) intron retention (Cartegni et al., 2002; Black, 2003).

that can either be skipped or be included in the mRNA; (2) the alternative 5′ splice site; (3) the alternative 3′ splice site; (4) the complex splicing pattern, when only one of the two mutually exclusive cassette exons is included in the mRNA; and (5) intron retention. Different types of alternative splicing can be combined to generate multiple mRNA isoforms from a single gene (Cartegni et al., 2002; Black, 2003). Besides that, different 5′ and 3′ ends of mRNA also can be generated by the use of alternative promoters in different exons and alternative polyadenylation signals.

Splicing Enhancers and Silencers

Exonic and intronic splicing enhancers (ESEs and ISEs) and silencers (ESSs and ISSs) are *cis*-regulatory elements that in most cases serve as binding sites for splicing factors, which stimulate or repress, respectively, splice-site usage or create an RNA secondary structure that affects splice-site recognition (Maniatis et al., 2002; Black, 2003). Frequently, enhancers and silencers overlap to form a composite regulatory element. ESEs are recognized by SR proteins, so far the best-studied splicing factors, which contain one or two RNA-binding domains and an arginine-serine (RS)–rich domain. The latter is thought to stimulate spliceosome assembly by recruiting U1 snRNP and U2AF via protein–protein interactions (Cartegni et al., 2002). However, recently it was also shown to interact directly with the pre-mRNA branch point (Shen et al., 2004). Moreover, SR proteins also can modulate splice-site recognition by competing with negative regulatory factors, but some may act as splicing repressors (Zhu et al., 2001; Cartegni et al., 2002; Maniatis et al., 2002).

ESEs have been identified through the analysis of disease-associated sequence variations (Pagani et al., 2004), by the in vivo or in vitro SELEX (systematic evolution of ligands by exponential enrichment) technique, which is based on the selection of sequences that promote splicing from a pool of random sequences (Cartegni et al., 2002), and by computational approaches based on statistical analysis of exon–intron and splice-site composition (Fairbrother et al., 2002). ESE consensus sequences have been derived from known SR protein binding sites, and the score matrices for individual SR proteins were calculated according to nucleotide frequencies at each position (Liu et al., 1998, 2000). SNPs that affect SR protein-binding scores in some cases have been shown to change splicing efficiency (Liu et al., 2001).

ESSs and ISSs are recognized by hnRNPs, a large group of diverse pre-mRNA binding proteins. Their mechanism of action is less well understood than that of enhancers. In some cases, enhancers and silencers partially overlap, and then the splicing efficiency is determined by the antagonistic action of hnRNPs and SR proteins. hnRNPs bound to ISSs flanking an alternative exon also may cause the exon to loop out, resulting in skipping of the exon. Alternatively, inhibitory factors bound to ESS may polymerize along the exon and displace the ESE-bound SR proteins (Zhu et al., 2001; Cartegni et al., 2002; Maniatis et al., 2002).

Intronic regulatory sequences often are within 50 bp of the splice sites. However, some regulatory elements are found thousands of bases from the splice sites. Some ISEs also are recognized by SR proteins, whereas others bind SR nonrelated proteins (Black, 2003; Pagani et al., 2004).

Regulation of Splice Site Selection

The selection of alternative splice sites can be regulated in different manners: tissue specificity, developmental stage, physiological processes, sex determination, and in response to various stress factors. A number of signals, including stimulation of receptors by growth factors, cytokines, or hormones; depolarization; rising intracellular Ca^{2+} levels; and cellular stresses like heat shock and change in pH, have been shown to induce changes in the selection of splice sites (Scotet et al., 1998; Stamm, 2002; Faustino et al., 2003). Splice-site selection in stress responses or in varying physiological environments can be regulated through changing the activity or composition of general

splicing factors by regulating transcription, by phosphorylation, or by changing their intracellular localization (Stamm, 2002). For example, insulin stimulates the inclusion of the B2 exon of protein kinase C (PKC), resulting in production of the PKCB2 isozyme. The binding of insulin to its receptor activates the phosphatidylinositol 3-kinase signaling pathway, leading to phosphorylation of the SR protein SRp40, which then promotes incorporation of the B2 exon (Patel et al., 2004). CD45, a transmembrane protein tyrosine phosphatase, is alternatively spliced in response to T-cell activation. The switch in CD45 splicing has been shown to be mediated by changing the expression level of several SR proteins, including SC35 (ten Dam et al., 2000; Wang et al., 2001a). In addition to general splicing factors present in multiple tissues, several cell-type specific splicing factors that regulate the splicing of specific subsets of genes have been identified. For example, the *Drosophila* halfpint protein (Hfp) regulates a number of genes required for oogenesis (Van Buskirk et al., 2002). So far, tissue-specific splicing factors in mammals predominantly have been found in the nervous system. One of the best-characterized examples is NOVA1, a neuron-specific RNA-binding protein that has been shown to stimulate the incorporation of exon E9 of *GABA(A)Rgamma2* by binding to an intronic splicing enhancer (Dredge et al., 2001, 2003).

METHODS FOR IDENTIFICATION AND ANALYSIS OF ALTERNATIVE SPLICE VARIANTS

So far, the most reliable technique for the identification and confirmation of alternative splice variants has been the sequencing of full-length cDNA. Various RT-PCR-, RACE (rapid amplification of cDNA ends)–, and Northern blot–based techniques are also widely used for the identification of alternative splice events. However, the power of these approaches for the characterization of alternative splicing patterns across all tissues, developmental stages, or disease conditions is limited. Recently, several approaches for the genome-wide analysis of AS have emerged, including various bioinformatic approaches based on the alignment of mRNAs and expressed sequence tags (ESTs) to the corresponding genomic sequence (Kan et al., 2001; Modrek et al., 2001, 2002; Xie et al., 2002; Xu et al., 2002), exon junction oligonucleotide microarrays (Johnson et al., 2003), AS profiling using a bead-based fiber-optic microarray platform (Yeakley et al., 2002), and polymerase colony (polony) technology (Mitra et al., 2003; Zhu et al., 2003).

Bioinformatic Approaches

More than 5.49 million ESTs derived from more than 6,000 human cancerous- and normal-tissue cDNA libraries have been deposited in the dbEST (http://www.ncbi.nlm.nih.gov/dbEST/) as of this writing (dbEST release 050704, May 7, 2004). These sequences provide a broad sample of mRNA diversity. However, it should be taken into account that genes with lower expression have a coverage bias and that all human tissues and developmental stages are not equally represented, with ESTs from brain and cell line libraries the most abundant. It is also a question whether alternative splice isoforms found in cell lines are functionally representative of those in the human organism. Nevertheless, EST data sets make a valuable source of information and, together with the available sequence of the human genome, have enabled the development of various bioinformatic approaches for the identification of alternative splice variants. Currently, bioinformatics is the only approach capable of identifying all types of novel AS forms on a genomewide scale. Most studies rely on the alignment of EST and mRNA sequences to each other (Wang et al., 2003) or to the genomic sequence (Kan et al., 2001; Modrek et al., 2001; Xie et al., 2002; Thanaraj et al., 2004). In the first step, clusters of ESTs corresponding to the same genomic locus are obtained either by a BLAST search of genomic (Kan et al., 2001) or RefSeq mRNA (Wang et al., 2003) sequences through dbEST or by using UNIGENE clusters (Modrek et al., 2001). Then the mRNA/ESTs themselves can be compared or they can be aligned to the genomic sequence using BLAST or specially designed software tools, such as *sim4* (http://www.hgmp.mrc.ac.uk/Registered/Webapp/sim4/) or Spidey (http://www.ncbi.nlm.nih.gov/IEB/Research/Ostell/Spidey/spideydoc.html), for aligning spliced cDNA sequences to genomic sequences. The alternative splicing events are inferred from the analyses of these alignments. Defining exon–intron boundaries is not a trivial task. Commonly, the alignment of two adjacent exons may overlap by 20–30 bp, and the true exon boundary may lie anywhere within this region. *Sim4* and Spidey use the score matrices of splice-site motifs to position exon boundaries. However, sometimes the inferred splice sites may deviate slightly from the correct sites, which may result in a false-positive prediction of 5′ or 3′ alternative

splice sites. Other problems that may cause errors in automated splice variant identification are coverage bias toward the 3′ end of transcripts, EST sequencing errors, contamination of EST databases with genomic or incompletely spliced mRNA sequences that may result in false predictions of intron retention events, and the presence of pseudogenes and paralogous genes in the human genome that may cause problems in EST clustering (Modrek et al., 2002; Johnson et al., 2003).

Nevertheless, bioinformatics-based studies have identified a vast number of alternatively spliced transcript variants and have estimated that 35%–59% of human genes are alternatively spliced (Modrek et al., 2002). However, the mRNA isoforms identified by automated analysis of ESTs should be taken only as a prediction and require further verification. So far, expression of the predicted splice variants has been experimentally verified only in a few studies (Brett et al., 2000; Wang et al., 2003). In these cases, RT-PCR with primers flanking the predicted alternative splice sites was used to analyze the expression of a small test set of transcripts. The low throughput of the method seems to be the bottleneck for experimental validation. Therefore, the development of high-throughput methods for the analysis of splice variants is needed to assess the specificity and sensitivity of the bioinformatics-based studies.

Recently, several large databases of alternative splicing have been developed (Huang et al., 2003; Thanaraj et al., 2004). In ProSplicer, alternative splicing forms are predicted by aligning not only ESTs and mRNAs but also protein sequences against the genomic sequence. In addition to the predicted exon–intron structures, it provides EST tissue information and links to OMIM, GO, and HUGO databases (Huang et al., 2003). A database developed by the Alternative Splicing Database (ASD) consortium contains data on AS events not only in humans, but also in mice, *Drosophila*, and other model organisms, allowing further validation of the identified splice variants by checking the preservation of the splice events across species. It also provides information about the characteristics of splice signals and provides links to GO classification and expression data (Thanaraj et al., 2004). The Extended Alternatively Spliced EST Database (EASED) project establishes a comprehensive database of alternative splice forms for nine eukaryotic organisms, including humans. The strength of this database is the provision of useful and detailed information (for humans only), for example, tissue type, developmental stage, disease notation, AS profile and classification of splice events, and the possibility of combining query parameters in order to filter out sequences of particular interest (Pospisil et al., 2004).

Exon Junction Microarrays

Johnson et al. (2003) demonstrated the use of oligonucleotide microarrays to monitor alternative exon usage. In this study, microarrays containing ~125,000 probes positioned at all known exon–exon junctions of 10,000 genes were hybridized to RNA samples derived from 52 different tissues and cell lines in order to identify tissue-specific splicing differences. The theoretical intensity of each probe was modeled as a function of the intensity of the probe response and tissue-specific expression level, and deviation between the observed and modeled intensities was used to predict an alternative splicing event. This approach proved to be useful for monitoring the expression of known splice variants and revealing novel exon-skipping events across different tissues. However, it generally is not capable of identifying novel 5′ or 3′ splice sites, of specifying the sequence of a novel splice variant, or of determining the combinatorial patterns of exon inclusion/skipping in the same transcript.

Bead-Based Fiber-Optic Microarrays

A novel approach that combines a fiber-optic microarray platform with a technique called RASL (RNA-mediated annealing, selection, and ligation) for a large-scale analysis of alternative splice variants was developed by Yeakley et al. (2002). In the RASL technique, the mRNA samples to be analyzed are annealed to multiple pairs of oligos complementary to exonic sequences flanking the presumed splice junctions. If the predicted splice variant is present, the corresponding two oligos, both carrying universal primer sites, are ligated enzymatically and amplified by PCR. The fluorescently labeled PCR products are then applied to a bead-based microarray platform that allows the quantity of each PCR product to be measured. Up to 100 splice junctions can be analyzed in a multiplexed fashion, and differences between mRNA samples in splicing efficiency and pattern can be measured.

This approach turned out to be reliable and extremely sensitive, allowing the detection of highly expressed mRNAs from less than 10 cells, which makes it particularly attractive for analyzing gene expression and splicing patterns in microdis-

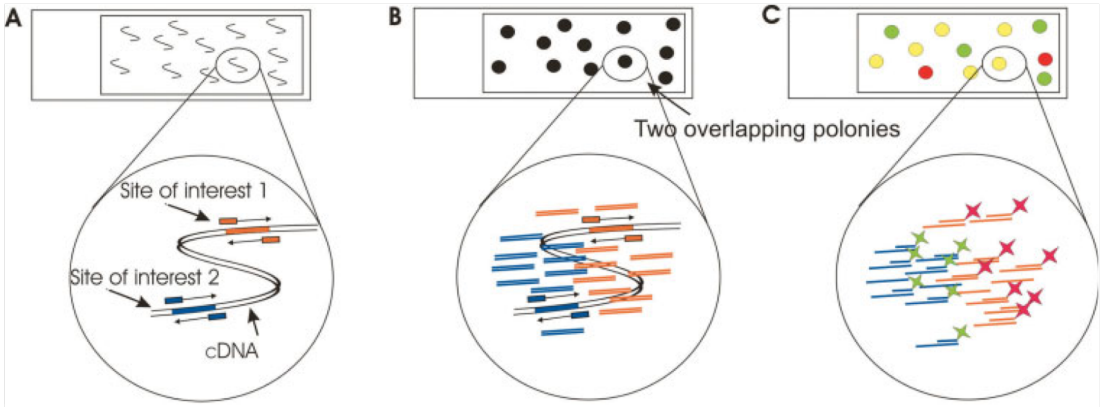


Figure 3. Polymerase colony (polony) technology. (A) Polymerization of diluted cDNA into a polyacrylamide gel containing all necessary reagents for PCR on a glass microscope slide. (B) Amplification of a single cDNA molecule with two primer pairs flanking the exons or the exon junctions of interest. Amplification products accumulate around the template, forming two overlapping polonies. One primer of each set was modified so that one strand of the amplified cDNA fragment was attached into the gel. All unattached strands were degraded and washed away after the amplification. (C) Polonies were visualized by hybridizing with exon-specific probes and scanning of slides. Green and red polonies represent transcripts with only one of the exon–exon junctions of interest; yellow polonies represent transcripts containing both exon–exon junctions of interest [adapted from Mitra et al. (2003) and Zhu et al. (2003)].

sected tissues. However, similarly to the exon junction microarrays, it requires knowledge of exon–intron structures and therefore is not applicable for searching for novel transcript variants or for the analysis of combinatorial splicing patterns.

Polymerase Colony (Polony) Technology

In the polony approach (Fig. 3), a small amount of cDNA or DNA template is polymerized into an acrylamide gel containing all the reagents necessary for PCR (Mitra et al., 2003; Zhu et al., 2003). In-gel PCR is performed on a glass microscope slide. Each template gives rise to a polony, an individual colony of amplification products. Amplification can be carried out simultaneously with several sets of primers flanking the splice sites of interest, thus generating overlapping polonies amplified from a different region of the same DNA molecule. One primer in each pair is modified in order to immobilize one strand of the amplified DNA on the acrylamide gel. This strand serves as a template for hybridization with fluorescently labeled exon-specific probes. The fluorescence is then recorded by a microarray scanner, and the images obtained are analyzed (merged) to determine the frequency of exon usage and the combinatorial patterns of exon inclusion and skipping and to quantify the expression of individual mRNA isoforms.

Currently, this is the only technique that allows multiple alternative splicing events in the same mRNA molecule to be monitored. Moreover, AS profiling can be combined with single-base extension (SBE) reactions in order to correlate SNPs with changes in the splicing pattern or efficiency (Zhu et al., 2003).

CHANGES IN PRE-mRNA SPLICING IN CANCER

A simple PubMed search with the keywords "splicing defects" or "aberrant splicing" and "cancer" found up to 100 genes whose pre-mRNA splicing is altered in various types of cancer. Both constitutive and alternative splicing may be changed, resulting in the production of novel, unnatural mRNA isoforms or in changes to the ratio or tissue specificity of natural mRNAs. Splicing defects have been observed in genes implicated in both the susceptibility and the progression of cancer. Several recent bioinformatic studies have revealed a vast number of potentially cancer-specific or cancer-associated splice variants. Wang et al. (2003) compared 11,014 mRNAs from the RefSeq database with about 3.47 million ESTs. That resulted in the identification of 26,258 alternative splice variants, of which 845 (3.2%) were significantly ($P < 0.05$) associated with cancer. The expression of 76 predicted splice variants was analyzed in paired specimens of tumor and adjacent tissue by RT-PCR, finding 72% of the expected products to be detectable and 59% to have a cancer-associated expression pattern. The analysis of 2 million ESTs in another genomewide survey of splicing aberrations identified 316 human genes that have cancer-specific splice variants (Xu et al., 2003). Hui et al. (2004) identified 383 potentially tumor-associated splice variants by aligning ESTs

with the genomic sequence of 4,322 genes. The frequencies of the identified alternative splice variants were compared between EST libraries derived from cancerous tissues with their corresponding normal counterparts and normalizing them against the expression level of each gene.

This raises a number of questions. What molecular mechanisms are responsible for changes in the recognition and usage of splice sites? Do splicing defects contribute to tumor development or progression? How frequent are splicing defects in human tumors? Is there a category of cancers in which splicing defects predominate, similar to the microsatellite instability or chromosome instability phenotype? Could the increase in splicing aberrations contribute to genetic instability? Could the protein isoforms generated as a result of splicing aberrations be recognized by a host's immune system and could they be used as targets for cancer immunotherapy?

Mechanisms of Splicing Defects

Generally, the mechanisms leading to splicing defects in cancer are poorly understood. It has been shown in individual cases that inherited or somatic mutations in *cis*-regulatory elements, as well as oncogenic signaling and variations in the composition, concentration, localization, and activity of *trans*-acting regulatory factors may result in the changes in splice-site recognition and usage.

Cis-Regulatory Elements

At least 15% of all disease-associated point mutations result in splicing defects, indicating that sequence variations in *cis*-regulatory elements could be responsible for a substantial number of missplicing events (Krawczak et al., 1992). However, whereas germ-line mutations resulting in splicing defects have been found in a number of genes implicated in a predisposition to cancer, such as *BRCA1*, *BRCA2*, *CDKN2A*, and *APC* (Liu et al., 2001; Charames et al., 2002; Agata et al., 2003; Rutter et al., 2003; Neklason et al., 2004), somatic mutations causing splicing defects rarely occur. Germ-line sequence variations in both splice sites and regulatory elements have been implicated in susceptibility to cancer. For example, an inherited nonsense mutation (Glu1694Ter) in exon 18 of the *BRCA1* gene disrupts an ESE, the binding site for SR protein SF2/ASF, causing inappropriate skipping of the constitutive exon 18 (Liu et al., 2001). Germ-line mutations in *APC* result in familial adenomatous polyposis (FAP). Mutations are scattered throughout the gene, with the precise location determining the severity of the disease (Neklason et al., 2004). Recently two mutations that disrupt splice regulatory elements have been found. The insertion of a T just beyond the conserved GT dinucleotide in the splice donor site of intron 4 results in the skipping of exon 4 and leads to an attenuated form of FAP (Neklason et al., 2004). A G-to-A substitution at the splice acceptor site of intron 7 creates a cryptic splice site, causing a single nucleotide deletion at the beginning of exon 8. A single base frameshift results in APC truncation and is associated with the classical polyposis phenotype (Charames et al., 2002).

Loss of heterozygosity and germ-line and somatic mutations in the *NF1* gene are frequently found in neurofibromas. A systematic study of somatic *NF1* mutations by Serra et al. (2001) demonstrated that most point mutations resulted in splicing defects, including exon skipping and the usage of alternative 5′ and 3′ splice sites. To our knowledge, this is the only study demonstrating that somatic mutations lead to altered splice-site recognition.

Oncogenic Signaling

CD44 is a multifunctional cell-surface glycoprotein involved in cell proliferation, differentiation, adhesion, and migration. Multiple CD44 isoforms can be generated by including and combining up to 10 variant internal exons under normal physiological conditions and during tumorigenesis. Certain CD44 splice variants, in particular those containing variant exons v5, v6, and v7, are overexpressed in various tumors and have been shown to play a role in tumor cell invasion and metastasis (Cooper et al., 1995; Naor et al., 2002; Faustino et al., 2003). The studies by Konig et al. (1998) and Matter et al. (2000) demonstrated that inclusion of exon v5 is regulated by a composite exonic splice regulator encompassing enhancer and silencer elements. The activation of the oncogenic RAS signal transduction pathway stimulates inclusion of exon v5. Because inducible splicing does not require de novo protein synthesis, the regulation of splicing is presumably carried out by a post-translational modification or by controlling the intracellular localization of splicing factors. In fact, Sam68, a member of the STAR (signal transduction and activation of RNA) protein family, was recently shown to become phosphorylated on activation of the ERK MAP-kinase pathway and was shown to be required for the RAS-induced inclusion of exon v5 (Matter et al., 2002).

Expression of *Trans*-Regulatory Elements

Several studies have demonstrated specific alterations in the expression of splicing factors in cancer. The increased phosphorylation and the elevated mRNA expression levels of Tra26, YB-1, and classical SR proteins, including SC35 and ASF/SF2, have been observed in human ovarian cancer and in mouse models of breast cancer development. Furthermore, these alterations correlate with the increased complexity of the CD44 splicing pattern (Stickeler et al., 1999; Fischer et al., 2004). Another study demonstrated that the p210 BCR/ABL7 fusion protein, generated by the Philadelphia translocation and considered responsible for initiation of chronic myelogenous leukemia, stimulates the expression of multiple genes involved in pre-mRNA splicing (e.g., SRPK1, RNA Helicase 2/Gu, hnRNPA2/B1, DDX10, and SF3B). This correlates with changes in the splicing efficiency of beta1-integrin-responsive nonreceptor tyrosin kinase (Salesse et al., 2004).

Although somatic splicing mutations are rarely found in cancers, ruling out an increased mutation rate as the principal mechanism for splicing defects, the current figures may be misleading. Generally, splicing enhancers and silencers, in particular those deep in introns, are poorly conserved, and therefore their identification and the evaluation of the impact of a nucleotide substitution on splicing efficiency relies mostly on experimental approaches on a case-by-case basis. It seems likely that many mutations currently classified as neutral, missense, or nonsense may turn out to affect the splicing process. In this context, it would be interesting to see whether tumors with mismatch repair defects also have an increased rate of splicing defects.

If the expression or activity of a splicing factor were altered in a cancer cell, it presumably would affect splicing of a subset of genes. Therefore, it should be possible to classify cancers according to their splicing defects. So far, there have been no studies done that would enable evaluation of which splicing factors have altered expression and/or activity, how frequently such alterations occur in human cancers, and what and how many genes are affected. It would also be of interest to see whether splicing defects confer a particular phenotype or have prognostic significance.

Functional Significance of Splicing Defects

At the protein level, a switch in splicing pattern can lead, for example, to an in-frame insertion or deletion of functional domains, the replacement of N or C termini, and/or truncated protein isoforms because of a frameshift. If an alternative transcript contains a premature termination codon more than 55 nucleotides upstream of the last exon–exon junction, the mRNA will most likely be degraded by the nonsense-mediated mRNA decay (NMD) pathway (Cartegni et al., 2002). AS can also generate mRNA isoforms differing in their 5′ or 3′ UTRs that may affect their stability, translation efficiency, and localization (Mignone et al., 2002). Whether splicing defects contribute to tumor development and progression or are just a consequence of the reduced fidelity of splicing machinery in cancer is still a controversial issue. Genetic instability acquired by mismatch repair defects or chromosome instability is a prerequisite for the selection of more aggressive tumor cells. Clearly, splicing machinery defects that affect a large number of genes would be lethal at the cellular level, whereas alterations in individual genes or moderate changes in the regulation of splicing might serve as additional means of increasing proteome complexity in a cancer cell (Philips et al., 2000). In some instances, aberrant splicing appears to serve as an alternative mechanism for inactivation of tumor suppressors or contributes to the gain of function of proteins promoting tumor development. Examples of aberrantly spliced genes contributing to tumor development or progression are listed in Table 1.

Inactivation of Putative Tumor Suppressors *FHIT* and *TSG101*

FHIT, a putative tumor suppressor, is a member of the histidine triad gene family encoding a diadenosine triphosphate hydrolase that is involved in purine metabolism (Barnes et al., 1996). In addition to the wild-type transcript, a vast number of different aberrant *FHIT* transcripts have been detected in various human tumors, including gastric, cervical, thyroid, and testicular germ-cell tumors (McIver et al., 2000; Lee et al., 2001, 2002; Huiping et al., 2002; Kraggerud et al., 2002; Terry et al., 2002). These transcripts are generated by exon skipping, by the use of alternative 5′ and 3′ splice sites, and by the recognition of cryptic splice sites, resulting in insertions of intronic sequences. No mutations have been found in the splice sites (McIver et al., 2000; Huiping et al., 2002), indicating that alterations in *trans*-acting regulators may be responsible for altered splice-site recognition. So far, no alternative FHIT protein isoforms have been described (Huiping et al., 2002), making it

TABLE 1. Genes Whose Splicing Pattern or Efficiency is Altered in Cancer

Gene	Splice variant	Mechanism	Functional impact	References
BRCA1	Skipped exon 18	Germ-line mutation (Glu1694Ter) disrupting ESE	Breast cancer susceptibility	Liu et al., 2001
APC	Skipped exon 4	Insertion of T beyond 5′ splice site	Attenuated FAP	Neklason et al., 2004
	1-bp deletion result in frameshift	G>A substitution created a cryptic splice site	Classical FAP	Charames et al., 2002
NF1	Exon skipping, alternative 5′ and 3′ splice sites	Various somatic mutations	Inactivation of tumor suppressor; neuroblastoma progression	Serra et al., 2001
CD44	Inclusion of exons v8–v10		Acquisition of metastatic potential	Shibuya et al., 1998
	Inclusion of exons v2–v10		Acquisition of metastatic potential	Barbour et al., 2003
	Inclusion of exons v5	Oncogene-regulated hnRNP A1 interaction to regulatory *cis*-acting splice elements	Tumor progression	Konig et al., 1998; Matter et al., 2000
FHIT	Exon skipping, alternative 5′ and 3′ splice sites	Possible alterations of *trans*-acting regulators	Inactivation of putative tumor suppressor by decreasing concentration of the functional mRNA	Huiping et al., 2002
TSG101	Exon skipping		Loss of candidate tumor-suppressor gene activity	Sun et al., 1997; McIver et al., 2000
RON (MSTIR)	Skipping of exon 11	Presence of a noncanonical splice-site consensus sequence	Acquisition of metastatic potential	Collesi et al., 1996
	Skipping of exons 5 and 6		Induction of cell scattering, transformation, and tumor growth in vivo	Zhou et al., 2003
HER2 (ERBB2)	Intron retention		Inhibition of growth factor–mediated tumor cell proliferation	Aigner et al., 2001
Fibronectin *(FN1)*	Insertion of a cryptic exon		Neovascularization/angiogenesis	Ebbinghaus et al., 2004
BIN1	Inappropriate expression of exon 12A		Abolished ability of Bin1 to inhibit malignant transformation by c-Myc or adenovirus EIA	Ge et al., 1999
MDM2	Alternative 5′ and 3′ splice sites	Aberrant splicing at direct repeat sequences	Loss of p53 binding domain. Activation of putative oncogene — promotion of p53-independent cell growth, inhibition of apoptosis	Lukas et al., 2001; Bartel et al., 2002
FAS (TNFRSF6)	Retention of intron 5		Impairment of Fas-induced apoptosis	van Doorn et al., 2002

unlikely that alternatively spliced products affect normal FHIT function in a dominant-negative manner. The splicing defects correlate with reduced FHIT protein production in most cases (McIver et al., 2000; Lee et al., 2001; Kraggerud et al., 2002). This suggests that deregulation of splicing could serve as an alternative mechanism for inactivation of the tumor suppressor by decreasing the level of functional mRNA. However, in several studies, some aberrant FHIT transcripts also were detected in normal tissues, making their implication for tumorigenesis uncertain (Gayther et al., 1997; Wang et al., 1999, 2000). This controversy presumably originated from the tissue distribution and functions of most of *FHIT* transcript variants not being characterized. It is therefore difficult to distinguish which variants represent tissue-specific alternative splicing and which represent splicing aberrations. It is also possible that aberrant transcript variants are produced because spliceosomal errors occur in both cancerous and normal tissue, but that they occur with increased frequency in cancer.

Another tumor suppressor, *TSG101,* whose expression appears to be important for the maintenance of genomic stability and cell-cycle regulation, is also a target of splicing defects and shows splicing abnormalities very similar to FHIT that are detectable in a wide range of tumors, including breast, ovary, prostate, thyroid, and cervical cancers and AML (Gayther et al., 1997; Sun et al., 1997; Klaes et al., 1999; Lin et al., 2000; McIver et al., 2000). Similarly, some transcript variants are also detectable in normal tissue, but the complexity and frequency of the aberrant splice variants is increased in tumors, showing a progressive loss of splicing fidelity during the malignant transformation (Klaes et al., 1999; Lin et al., 2000).

Aberrant *MDM2* Splicing Correlates with TP53 Status and Overall Survival

MDM2 is a putative oncogene that has been implicated in various cellular processes, including regulation of cell growth and apoptosis. MDM2 and TP53 form an autoregulatory loop in which the transcription of the *MDM2* gene is up-regulated by TP53, whereas the binding of MDM2 to TP53 inhibits transactivation by TP53 and promotes its degradation by proteasomes (Bartel et al., 2002, 2004). More than 40 *MDM2* transcript variants have been identified, but the pattern of alternative splicing and the number of transcripts differ between normal and cancerous tissues. The encoded proteins are likely to be functionally different. Thus, for example, MDM2-b, one of the splice forms lacking a TP53-binding domain has been shown to promote TP53-independent cell growth, to inhibit apoptosis, and to up-regulate the RelA subunit of NFkappaB (Steinman et al., 2004). Some transcripts generated by the use of cryptic splice sites and the skipping of exons were detectable exclusively in invasive breast cancer, not in normal breast tissue. No mutations that could create these cryptic splice sites were found in these tumors, but sequence analysis demonstrated that splicing occurred between direct repeat sequences. Moreover, the expression of the aberrant transcripts was correlated with *TP53* overexpression, *TP53* mutations, and shorter overall survival of breast cancer patients, suggesting that aberrant splice forms may play a significant role in tumorigenesis and might serve as a prognostic marker for breast cancer (Lukas et al., 2001).

Generation of Constitutively Active RON Tyrosine Kinase Isoforms

Another example of how alternative splicing contributes to tumor growth and the acquisition of an invasive phenotype is RON (MST1R according to the HUGO nomenclature). RON is a tyrosine kinase receptor for macrophage-stimulating protein (MSP). Wild-type RON is a heterodimeric protein composed of α and β subunits derived from proteolytic cleavage of a common precursor. The binding of MSP to RON triggers an intracellular signaling cascade eliciting cell dissociation, motility, and invasiveness (Comoglio et al., 1996). Three alternatively spliced RON transcripts, encoding 165-, 160-, and 155-kDa RON isoforms, have been identified. RONΔ165 was found in a gastric cancer cell line (Collesi et al., 1996), whereas all three isoforms were detected in primary colorectal carcinomas (Zhou et al., 2003). All three RON variants are constitutively active but differ in their biochemical and biological properties (Zhou et al., 2003). RONΔ165 is generated by skipping a 147-bp exon, resulting in a 49-amino-acid deletion from the extracellular domain, which prevents the precursor from undergoing proteolytic cleavage. The unprocessed precursor is retained in the cytoplasm, and the tyrosine kinase is constitutively activated by oligomerization and autophosphorylation. Cells expressing RONΔ165 acquired invasive properties in vitro (Collesi et al., 1996). RONΔ155 also is not cleaved and is present as a single-chain precursor. RONΔ160 is produced by skipping exons 5 and 6 (109 amino acids from the extracellular domain),

but this does not interfere with the cleavage of the precursor, and the heterodimeric 160-kDa protein is located on the cell surface and binds MSP. Both the 160- and 155-kDa isoforms induce cell scattering, transformation, and tumor growth in vivo (Zhou et al., 2003).

IMMUNOLOGICAL RECOGNITION OF ABERRANTLY SPLICED GENE PRODUCTS IN CANCER

Recently, it has been shown that frameshift-mutation-derived peptides are recognized by tumor-infiltrating lymphocytes in MSI+ colon cancer patients, and it has been suggested that these are very attractive targets for cancer vaccines (Saeterdal et al., 2001; Ishikawa et al., 2003). Similarly, peptides derived form aberrantly spliced transcripts could be recognized by the immune system and serve as targets for cancer immunotherapy. Novel exon–exon junctions generated by exon skipping, intron retention (in-frame and out-of-frame translations), and the inclusion of cryptic exons may result in structural alterations of a protein and represent novel T- and B-cell epitopes. In fact, several proteins encoded by alternatively or aberrantly spliced genes have been shown to induce an immune response in cancer patients (Table 2).

Peptides Derived from Aberrantly Spliced Transcripts as CTL Targets

Yannik Guilloux et al. (1996) identified a peptide derived from an intronic sequence of the N-acetylglucosaminyltransferase V (*GnT-V*/*MGAT5*) gene as a novel tumor antigen, NA17-A, that is recognized by cytotoxic T lymphocytes (CTLs) on HLA-A2 melanomas. The mRNA encoding the NA17-A antigen is generated by the use of an alternative promoter in an intron and is translated in a different reading frame than *GnT-V*. The aberrant transcript was expressed at a relatively high level in 50% of HLA-A2 melanomas, whereas only trace levels were detectable in normal tissues (Guilloux et al., 1996). Peptide-pulsed and apoptotic body-loaded dendritic cells stimulate NA17-A-specific CTL activation, demonstrating that NA17-A could be a useful target for cancer immunotherapy (Labarriere et al., 1998, 2002).

The tyrosinase-related protein (TRP) 2-INT2 isoform represents another case of intron retention that results in the generation of a melanoma-restricted T-cell epitope. *TRP2* itself encodes melanoma differentiation antigen, which has its own antigenic epitope expressed in melanoma and melanocytes (Wang et al., 1996). The TRP2-INT2 isoform contains exons 1–4, with retention of intron 2 and a part of intron 4. The translation of intron 2 in the same reading frame of the fully spliced TRP2 mRNA introduces a stop codon at the 3′ end of the intron. In contrast to TRP2, this isoform is highly expressed only in melanoma cells, not in normal melanocytes, suggesting that the splicing pattern of TRP2 is altered in melanoma. The peptide encoded by the intronic sequence is recognized by CTLs when exposed on HLA-A*68011 (Lupetti et al., 1998). Recently another TRP2 isoform, TRP2-6b, generated by the insertion of two cryptic exons (6b and 6c) between exon 6 and 7, was identified. It is translated in-frame with the wt TRP2 and contains an insertion of 33 amino acids. Two peptides encoded by exon 6b were recognized by melanoma-specific HLA-A2 restricted tumor-infiltrated lymphocytes (Khong et al., 2002).

Autoantibody Responses Against Alternatively Spliced Gene Products

The SEREX (serological identification of antigens by recombinant expression cloning) approach is based on the screening of cDNA expression libraries with cancer patients' sera, allowing a systematic search for genes whose products have elicited autoantibody production in cancer patients (Sahin et al., 1995, 2001). This technique has been applied to various types of tumors, resulting in the identification of approximately 1,450 potential tumor antigens. So far, the molecular mechanisms underlying the immunogenicity of the majority of SEREX-identified antigens are unknown, but it is assumed that the antibody responses may be associated with a cancer-associated expression pattern (cancer-testis and differentiation antigens), mutations, and overexpression (Türeci et al., 1999; Pfreundschuh, 2000). A large number of these genes have multiple transcript variants generated by alternative promoter and splice-site usage, raising the possibility that alterations in splicing might be associated with their immunogenicity. In fact, alterations in splicing pattern or efficiency and the use of alternative promoters in cancer have been demonstrated for several genes, including *ELAV* (Hu antigen D), *TACC1*, *CML66*, *CTAGE*, and *XAGE1* (Behrends et al., 2002; Line et al., 2002a; Usener et al., 2003; Ali Eldib et al., 2004; Yan et al., 2004).

We previously identified a microtubule-associated protein, TACC1, as a potential tumor antigen by applying the SEREX technique to gastric cancer (Line et al., 2002a, 2002b). The serum-reactive cDNA clone represented a novel splice variant of

TABLE 2. Splice Variants Inducing Cellular or Humoral Immune Responses in Cancer Patients

Gene	Splicing pattern	Tumor type	Expression pattern	T- or B-cell response	Epitope	References
XAGE1 (GAGED2)	4 transcript variants: 1a, 1b, 1c, 1d	Lung cancer	Cancer — testes	Autoantibody responses to XAGE-1b	ND	Ali Eldib et al., 2004
				Autoantibody responses to XAGE-1c (L552S)	N-terminal peptide unique to 1c isoform	Wang et al., 2001b
TACC1	10 transcript variants (*TACC1-A–TACC1-J*)	Gastric cancer	*TACC1-F*: normal brain and gastric cancer; *TACC1-D*: gastric cancer, trace amounts in normal brain and stomach	Autoantibody responses to TACC1	ND	Line et al., 2002a, 2002b
GnT-V *(MGAT5)*	2 transcript variants generated by usage of alternative promoter	Malignant melanoma	*NA17-A* encoding isoform expressed at high levels in melanomas, trace levels detectable in normal tissues	CTL responses to antigen NA17-A	Peptide generated from retained intron	Guilloux et al., 1996
ELAV (Hu antigen D)	4 transcript variants: *HuD1–HuD4*	Childhood neuroblastoma	*HUD1* expression restricted to CNS and neuroectodermal tumors	Autoantibody responses to HUD1	ND	Behrends et al., 2002
CML66	2 transcript variants: *CML66-L* and *–S TRP 2-INT2*	CML	*CML66-L*: cancer—testes; *CML66-S*: testes, barely in cancer	Autoantibody responses to CML66-L	5' region of protein isoform CML66-L	Yan et al., 2004
TRP2 (TYRP1)	*TRP2-INT2* (retention of intron 2)	Malignant melanoma	Melanoma—specific	CTL responses to TRP-2-INT2 and TRP-2-6b	Peptides encoded by intron 2	Wang et al., 1996; Khong et al., 2002
	TRP-2-6b (insertion of novel exons 6b and 6c)		?		2 peptides encoded by exon 6b	Khong et al., 2002
CTAGE-1	3 transcript variants: *cTAGE-1, -1B, -1C*	Cutaneous T cell lymphoma	*CTAGE*-1 and -1B: cancer—testes	Autoantibody responses to cTAGE-1	Identified for cTAGE-1	Usener et al., 2003
TAG	4 transcript variants: TAG-1, *TAG-2a, TAG-2b, TAG-2c*	Melanoma	Cancer—testes	CTL responses to TAG-derived peptide	Peptide encoded by junction of the first two exons	Hogan et al., 2004
ING1	4 transcript variants: A, B, C, D	Breast cancer	A — ubiquitous B — differential, some tumors; C — testes, weakly in brain	Autoantibody responses to ING1	ND	Jager et al., 1999

ND, not determined.

TACC1 generated by insertion of a 36-bp exon. Analysis of the *TACC1* mRNA structure and expression pattern revealed 10 novel transcript variants and demonstrated that the regulation of alternative splicing of *TACC1* pre-mRNA is altered in gastric cancer, resulting in overexpression of two splice variants, *TACC1-F* and *-D*. The two transcript variants share the same 36-bp exon, which is not included in any other transcript. *TACC1-F* was predominantly expressed in normal brain and at comparable levels in 50% of gastric cancer specimens, which was 50- to 1,000-fold greater than the expression in paired adjacent tissue and in other normal tissue. *TACC1-D* was expressed in 52% of gastric tumors, but only trace levels were detectable in adjacent gastric mucosa and normal brain (unpublished data). It seems likely that the production of *TACC1-F* and *-D* isoforms in cancerous tissues may underlie the immunogenicity of TACC1. However, which region of the protein is recognized by autoantibodies and whether the isoform-specific peptides are capable of inducing CTL responses remain to be determined.

XAGE-1 (G antigen, family D) belongs to the family of cancer-testis antigens. Four transcript variants are generated by alternative splicing and alternate transcription initiation sites. Recently, two of these isoforms—XAGE-1b and 1c, were detected by SEREX analysis of lung cancer (Ali Eldib et al., 2004). In this study, XAGE-1b was shown to be the dominant transcript in testis and lung cancer, and it was recognized by sera from 8 of 32 cancer patients but not by healthy controls. In a study by Wang et al. (2001b), another XAGE-1 isoform, L552S (corresponding to XAGE-1c), was shown to be overexpressed and immunogenic in lung cancer. Epitope mapping showed that the N-terminal region, which is unique in the XAGE-1c isoform, was recognized by autoantibodies in at least one patient.

Six antigens encoded by novel splice variants were identified by SEREX in chronic lymphocytic leukemia (CLL). One of the clones, KW-4, encodes a Kruppel-like zinc finger protein, ZNF286, missing the KRAB domain. Full-length ZNF286 is ubiquitously expressed, whereas the identified splice variant was not detected in 8 normal tissue samples but was expressed in all CLL cases tested (Krackhardt et al., 2002). However, it is not known whether the immune response is directed against the splice-variant-specific epitope.

Recently, six transcript variants of a microtubule-associated protein, kinectin, were found to induce antibody responses in patients with hepatocellular carcinoma (Wang et al., 2004). One of the variants showed cancer-associated overexpression, suggesting a putative association of the altered splicing efficiency with the immunogenicity.

We have described here how splicing defects may contribute to the generation of novel epitopes. However, it is possible that the alterations in splicing also could change the antigen repertoire of tumor cells by removing immunodominant epitopes, contributing to the immune escape of tumors. In fact, the differential splicing of RNA encoding a CTL epitope of EBV has been shown to reduce the functional presentation of the immunodominant EBV epitope (Kienzle et al., 2000).

In several instances described here (TRP2-INT2, HER2, FAS, etc.), missplicing events introduced premature termination codons. Thus, in theory, these transcripts should be targeted for degradation by the NMD pathway, a conserved proofreading mechanism in eukaryotic cells that protects them from the potentially detrimental effects of truncated proteins (Byers, 2002). Nevertheless, we have summarized the evidence that aberrant transcripts can be translated into functional proteins, generating novel cancer-specific antigenic epitopes recognized by both the cellular and humoral responses of the host's immune system. This raises an intriguing possibility: that the NMD mechanism by itself might be disturbed in cancers, resulting in the accumulation of aberrant transcripts that escape from NMD.

We have attempted to provide insight into the mechanisms and the significance of pre-mRNA splicing alterations in cancer. We have only mentioned a small number of genes whose splicing is altered in cancer. These alterations result either in reduction of the number of normal transcripts, thus contributing to inactivation of tumor suppressors, or in the production of novel protein isoforms conferring oncogenic or metastatic potential to the tumor cells. In some cases, the proteins generated from the aberrant transcripts can be recognized by the host's immune system, revealing their relevance as immunotherapeutic targets. However, a line of evidence suggests that splicing defects are an essential feature of cancer. New tools allowing systematic large-scale analyses of alternative splicing efficiency and patterns are emerging. Hopefully, this will lead to a better understanding of the nature and the role of splicing defects in cancer. We anticipate that the determination of splicing patterns in cancer will provide novel diagnostic and prognostic markers and may serve as a source

for the identification of novel targets for the immunotherapy of cancer. Furthermore, the splice variants that functionally contribute to tumor development or progression may be targeted by various molecular therapies that appear to be close to clinical application, which include targeting by small molecules of aberrant protein isoforms or splicing factors, inhibition or activation of specific splicing events by the use of various oligonucleotide-mediated therapies, and RNA-based reprogramming of alternative splicing (for a detailed review, see Garcia-Blanco et al., 2004).

REFERENCES

Agata S, De Nicolo A, Chieco-Bianchi L, D'Andrea E, Menin C, Montagna M. 2003. The BRCA2 sequence variant IVS19+1G→A A leads to an aberrant transcript lacking exon 19. Cancer Genet Cytogenet 141:175–176.

Aigner A, Juhl H, Malerczyk C, Tkybusch A, Benz CC, Czubayko F. 2001. Expression of a truncated 100 kDa HER2 splice variant acts as an endogenous inhibitor of tumour cell proliferation. Oncogene 20:2101–2111.

Ali Eldib AM, Ono T, Shimono M, Kaneko M, Nakagawa K, Tanaka R, Noguchi Y, Nakayama E. 2004. Immunoscreening of a cDNA library from a lung cancer cell line using autologous patient serum: identification of XAGE-1b as a dominant antigen and its immunogenicity in lung adenocarcinoma. Int J Cancer 108:558–563.

Barbour AP, Reeder JA, Walsh MD, Fawcett J, Antalis TM, Gotley DC. 2003. Expression of the CD44v2-10 isoform confers a metastatic phenotype: importance of the heparan sulfate attachment site CD44v3. Cancer Res 63:887–892.

Barnes LD, Garrison PN, Siprashvili Z, Guranowski A, Robinson AK, Ingram SW, Croce CM, Ohta M, Huebner K. 1996. Fhit, a putative tumor suppressor in humans, is a dinucleoside 5′,5′′-P1,P3-triphosphate hydrolase. Biochemistry 35:11529–11535.

Bartel F, Taubert H, Harris LC. 2002. Alternative and aberrant splicing of MDM2 mRNA in human cancer. Cancer Cell 2:9–15.

Bartel F, Harris LC, Wurl P, Taubert H. 2004. MDM2 and its splice variant messenger RNAs: expression in tumors and down-regulation using antisense oligonucleotides. Mol Cancer Res 2:29–35.

Behrends U, Jandl T, Golbeck A, Lechner B, Muller-Weihrich S, Schmid I, Till H, Berthold F, Voltz R, Mautner JM. 2002. Novel products of the HUD, HUC, NNP-1 and alpha-internexin genes identified by autologous antibody screening of a pediatric neuroblastoma library. Int J Cancer 100:669–677.

Black DL. 2003. Mechanisms of alternative pre-messenger RNA splicing. Annu Rev Biochem 72:291–336.

Brett D, Hanke J, Lehmann G, Haase S, Delbruck S, Krueger S, Reich J, Bork P. 2000. EST comparison indicates 38% of human mRNAs contain possible alternative splice forms. FEBS Lett 474:83–86.

Burset M, Seledtsov IA, Solovyev VV. 2000. Analysis of canonical and non-canonical splice sites in mammalian genomes. Nucleic Acids Res 28:4364–4375.

Byers PH. 2002. Killing the messenger: new insights into nonsense-mediated mRNA decay. J Clin Invest 109:3–6.

Cartegni L, Chew SL, Krainer AR. 2002. Listening to silence and understanding nonsense: exonic mutations that affect splicing. Nat Rev Genet 3:285–298.

Charames GS, Cheng H, Gilpin CA, Hunter AG, Berk T, Bapat B. 2002. A novel aberrant splice site mutation in the APC gene. J Med Genet 39:754–757.

Collesi C, Santoro MM, Gaudino G, Comoglio PM. 1996. A splicing variant of the RON transcript induces constitutive tyrosine kinase activity and an invasive phenotype. Mol Cell Biol 16: 5518–5526.

Comoglio PM, Boccaccio C. 1996. The HGF receptor family: unconventional signal transducers for invasive cell growth. Genes Cells 1:347–354.

Cooper DL, Dougherty GJ. 1995. To metastasize or not? Selection of CD44 splice sites. Nat Med 1:635–637.

Dredge BK, Darnell RB. 2003. Nova regulates GABA(A) receptor gamma2 alternative splicing via a distal downstream UCAU-rich intronic splicing enhancer. Mol Cell Biol 23:4687–4700.

Dredge BK, Polydorides AD, Darnell RB. 2001. The splice of life: alternative splicing and neurological disease. Nat Rev Neurosci 2:43–50.

Ebbinghaus C, Scheuermann J, Neri D, Elia G. 2004. Diagnostic and therapeutic applications of recombinant antibodies: targeting the extra-domain B of fibronectin, a marker of tumor angiogenesis. Curr Pharm Des 10:1537–1549.

Fairbrother WG, Yeh RF, Sharp PA, Burge CB. 2002. Predictive identification of exonic splicing enhancers in human genes. Science 297:1007–1013.

Faustino NA, Cooper TA. 2003. Pre-mRNA splicing and human disease. Genes Dev 17:419–437.

Fischer DC, Noack K, Runnebaum IB, Watermann DO, Kieback DG, Stamm S, Stickeler E. 2004. Expression of splicing factors in human ovarian cancer. Oncol Rep 11:1085–1090.

Garcia-Blanco MA, Baraniak AP, Lasda EL. 2004. Alternative splicing in disease and therapy. Nat Biotechnol 22:535–546.

Gayther SA, Barski P, Batley SJ, Li L, de Foy KA, Cohen SN, Ponder BA, Caldas C. 1997. Aberrant splicing of the TSG101 and FHIT genes occurs frequently in multiple malignancies and in normal tissues and mimics alterations previously described in tumours. Oncogene 15:2119–2126.

Ge K, DuHadaway J, Du W, Herlyn M, Rodeck U, Prendergast GC. 1999. Mechanism for elimination of a tumor suppressor: aberrant splicing of a brain-specific exon causes loss of function of Bin1 in melanoma. Proc Natl Acad Sci USA 96:9689–9694.

Guilloux Y, Lucas S, Brichard VG, Van Pel A, Viret C, De Plaen E, Brasseur F, Lethe B, Jotereau F, Boon T. 1996. A peptide recognized by human cytolytic T lymphocytes on HLA-A2 melanomas is encoded by an intron sequence of the N-acetylglucosaminyltransferase V gene. J Exp Med 183:1173–1183.

Hogan KT, Coppola MA, Gatlin CL, Thompson LW, Shabanowitz J, Hunt DF, Engelhard VH, Ross MM, Slingluff CL, Jr. 2004. Identification of novel and widely expressed cancer/testis gene isoforms that elicit spontaneous cytotoxic T-lymphocyte reactivity to melanoma. Cancer Res 64:1157–1163.

Huang HD, Horng JT, Lee CC, Liu BJ. 2003. ProSplicer: a database of putative alternative splicing information derived from protein, mRNA and expressed sequence tag sequence data. Genome Biol 4:R29.

Hui L, Zhang X, Wu X, Lin Z, Wang Q, Li Y, Hu G. 2004. Identification of alternatively spliced mRNA variants related to cancers by genome-wide ESTs alignment. Oncogene 23:3013–3023.

Huiping C, Kristjansdottir S, Bergthorsson JT, Jonasson JG, Magnusson J, Egilsson V, Ingvarsson S. 2002. High frequency of LOH, MSI and abnormal expression of FHIT in gastric cancer. Eur J Cancer 38:728–735.

Ishikawa T, Fujita T, Suzuki Y, Okabe S, Yuasa Y, Iwai T, Kawakami Y. 2003. Tumor-specific immunological recognition of frameshift-mutated peptides in colon cancer with microsatellite instability. Cancer Res 63:5564–5572.

Jager D, Stockert E, Scanlan MJ, Gure AO, Jager E, Knuth A, Old LJ, Chen YT. 1999. Cancer-testis antigens and ING1 tumor suppressor gene product are breast cancer antigens: characterization of tissue-specific ING1 transcripts and a homologue gene. Cancer Res 59:6197–6204.

Johnson JM, Castle J, Garrett-Engele P, Kan Z, Loerch PM, Armour CD, Santos R, Schadt EE, Stoughton R, Shoemaker DD. 2003. Genome-wide survey of human alternative pre-mRNA splicing with exon junction microarrays. Science 302:2141–2144.

Kan Z, Rouchka EC, Gish WR, States DJ. 2001. Gene structure prediction and alternative splicing analysis using genomically aligned ESTs. Genome Res 11:889–900.

Khong HT, Rosenberg SA. 2002. Pre-existing immunity to tyrosinase-related protein (TRP)-2, a new TRP-2 isoform, and the NY-ESO-1 melanoma antigen in a patient with a dramatic response to immunotherapy. J Immunol 168:951–956.

Kienzle N, Buck M, Silins SL, Burrows SR, Moss DJ, Winterhalter A, Brooks A, Khanna R. 2000. Differential splicing of antigen-encoding RNA reduces endogenous epitope presentation that regulates the expansion and cytotoxicity of T cells. J Immunol 165:1840–1846.

Klaes R, Kloor M, Willeke F, Melsheimer P, von Knebel DM, Ridder R. 1999. Significant increase of a specific variant TSG101 transcript during the progression of cervical neoplasia. Eur J Cancer 35:733–737.

Konig H, Ponta H, Herrlich P. 1998. Coupling of signal transduction to alternative pre-mRNA splicing by a composite splice regulator. EMBO J 17:2904–2913.

Krackhardt AM, Witzens M, Harig S, Hodi FS, Zauls AJ, Chessia M, Barrett P, Gribben JG. 2002. Identification of tumor-associated antigens in chronic lymphocytic leukemia by SEREX. Blood 100:2123–2131.

Kraggerud SM, Aman P, Holm R, Stenwig AE, Fossa SD, Nesland JM, Lothe RA. 2002. Alterations of the fragile histidine triad gene, FHIT, and its encoded products contribute to testicular germ cell tumorigenesis. Cancer Res 62:512–517.

Krawczak M, Reiss J, Cooper DN. 1992. The mutational spectrum of single base-pair substitutions in mRNA splice junctions of human genes: causes and consequences. Hum Genet 90:41–54.

Labarriere N, Pandolfino MC, Raingeard D, Le Guiner S, Diez E, Le Drean E, Dreno B, Jotereau F. 1998. Frequency and relative fraction of tumor antigen-specific T cells among lymphocytes from melanoma-invaded lymph nodes. Int J Cancer 78:209–215.

Labarriere N, Bretaudeau L, Gervois N, Bodinier M, Bougras G, Diez E, Lang F, Gregoire M, Jotereau F. 2002. Apoptotic body-loaded dendritic cells efficiently cross-prime cytotoxic T lymphocytes specific for NA17-A antigen but not for Melan-A/MART-1 antigen. Int J Cancer 101:280–286.

Lander ES, Linton LM, Birren B, Nusbaum C et al. 2001. Initial sequencing and analysis of the human genome. Nature 409:860–921.

Lee SH, Kim CJ, Park HK, Koh JW, Cho MH, Baek MJ, Lee MS. 2001. Characterization of aberrant FHIT transcripts in gastric adenocarcinomas. Exp Mol Med 33:124–130.

Lee SH, Kim HY, Kim TJ, Park HK, Kim WH, Woo KM, Cho MH. 2002. Aberrant splicing of FHIT transcripts in human gastric cancer cell lines. Res Commun Mol Pathol Pharmacol 112:39–49.

Lin SF, Lin PM, Liu TC, Chang JG, Sue YC, Chen TP. 2000. Clinical implications of aberrant TSG101 transcripts in acute myeloblastic leukemia. Leuk Lymphoma 36:463–466.

Line A, Slucka Z, Stengrevics A, Li G, Rees RC. 2002a. Altered splicing pattern of TACC1 mRNA in gastric cancer. Cancer Genet Cytogenet 139:78–83.

Line A, Stengrevics A, Slucka Z, Li G, Jankevics E, Rees RC. 2002b. Serological identification and expression analysis of gastric cancer-associated genes. Br J Cancer 86:1824–1830.

Liu HX, Zhang M, Krainer AR. 1998. Identification of functional exonic splicing enhancer motifs recognized by individual SR proteins. Genes Dev 12:1998–2012.

Liu HX, Chew SL, Cartegni L, Zhang MQ, Krainer AR. 2000. Exonic splicing enhancer motif recognized by human SC35 under splicing conditions. Mol Cell Biol 20:1063–1071.

Liu HX, Cartegni L, Zhang MQ, Krainer AR. 2001. A mechanism for exon skipping caused by nonsense or missense mutations in BRCA1 and other genes. Nat Genet 27:55–58.

Lukas J, Gao DQ, Keshmeshian M, Wen WH, Tsao-Wei D, Rosenberg S, Press MF. 2001. Alternative and aberrant messenger RNA splicing of the mdm2 oncogene in invasive breast cancer. Cancer Res 61:3212–3219.

Lupetti R, Pisarra P, Verrecchia A, Farina C, Nicolini G, Anichini A, Bordignon C, Sensi M, Parmiani G, Traversari C. 1998. Translation of a retained intron in tyrosinase-related protein (TRP) 2 mRNA generates a new cytotoxic T lymphocyte (CTL)-defined and shared human melanoma antigen not expressed in normal cells of the melanocytic lineage. J Exp Med 188:1005–1016.

Maniatis T, Tasic B. 2002. Alternative pre-mRNA splicing and proteome expansion in metazoans. Nature 418:236–243.

Matter N, Marx M, Weg-Remers S, Ponta H, Herrlich P, Konig H. 2000. Heterogeneous ribonucleoprotein A1 is part of an exon-specific splice-silencing complex controlled by oncogenic signaling pathways. J Biol Chem 275:35353–35360.

Matter N, Herrlich P, Konig H. 2002. Signal-dependent regulation of splicing via phosphorylation of Sam68. Nature 420:691–695.

McIver B, Grebe SK, Wang L, Hay ID, Yokomizo A, Liu W, Goellner JR, Grant CS, Smith DI, Eberhardt NL. 2000. FHIT and TSG101 in thyroid tumours: aberrant transcripts reflect rare abnormal RNA processing events of uncertain pathogenetic or clinical significance. Clin Endocrinol (Oxf) 52:749–757.

Mignone F, Gissi C, Liuni S, Pesole G. 2002. Untranslated regions of mRNAs. Genome Biol 3:REVIEWS0004-

Mitra RD, Butty VL, Shendure J, Williams BR, Housman DE, Church GM. 2003. Digital genotyping and haplotyping with polymerase colonies. Proc Natl Acad Sci USA 100:5926–5931.

Modrek B, Lee C. 2002. A genomic view of alternative splicing. Nat Genet 30:13–19.

Modrek B, Resch A, Grasso C, Lee C. 2001. Genome-wide detection of alternative splicing in expressed sequences of human genes. Nucleic Acids Res 29:2850–2859.

Naor D, Nedvetzki S, Golan I, Melnik L, Faitelson Y. 2002. CD44 in cancer. Crit Rev Clin Lab Sci 39:527–579.

Neklason DW, Solomon CH, Dalton AL, Kuwada SK, Burt RW. 2004. Intron 4 mutation in APC gene results in splice defect and attenuated FAP phenotype. Fam Cancer 3:35–40.

Pagani F, Baralle FE. 2004. Genomic variants in exons and introns: identifying the splicing spoilers. Nat Rev Genet 5:389–396.

Patel AA, Steitz JA. 2003. Splicing double: insights from the second spliceosome. Nat Rev Mol Cell Biol 4:960–970.

Patel NA, Apostolatos HS, Mebert K, Chalfant CE, Watson JE, Pillay TS, Sparks J, Cooper DR. 2004. Insulin regulates protein kinase CbetaII alternative splicing in multiple target tissues: development of a hormonally responsive heterologous minigene. Mol Endocrinol 18:899–911.

Pfreundschuh M. 2000. Exploitation of the B cell repertoire for the identification of human tumor antigens. Cancer Chemother Pharmacol 46Suppl:S3–S7.

Philips AV, Cooper TA. 2000. RNA processing and human disease. Cell Mol Life Sci 57:235–249.

Pospisil H, Herrmann A, Bortfeldt RH, Reich JG. 2004. EASED: Extended Alternatively Spliced EST Database. Nucleic Acids Res 32:D70–D74.

Rutter JL, Goldstein AM, Davila MR, Tucker MA, Struewing JP. 2003. CDKN2A point mutations D153spl(c.457G>T) and IVS2+1G>T result in aberrant splice products affecting both p16INK4a and p14ARF. Oncogene 22:4444–4448.

Saeterdal I, Bjorheim J, Lislerud K, Gjertsen MK, Bukholm IK, Olsen OC, Nesland JM, Eriksen JA, Moller M, Lindblom A, Gaudernack G. 2001. Frameshift-mutation-derived peptides as tumor-specific antigens in inherited and spontaneous colorectal cancer. Proc Natl Acad Sci USA 98:13255–13260.

Sahin U, Tureci O, Schmitt H, Cochlovius B, Johannes T, Schmits R, Stenner F, Luo G, Schobert I, Pfreundschuh M. 1995. Human neoplasms elicit multiple specific immune responses in the autologous host. Proc Natl Acad Sci USA 92: 11810–11813.

Sahin U, Li G, Tureci O, Pfreundschuh M. 2001. Recognition of human tumors: SEREX expression cloning to identify tumour antigens. In: Robins RA, Rees RC (eds). Cancer Immunology. Dordrecht/Boston/London: Kluwer Academic Publishers. p 45–57.

Salesse S, Dylla SJ, Verfaillie CM. 2004. p210BCR/ABL-induced alteration of pre-mRNA splicing in primary human CD34+ hematopoietic progenitor cells. Leukemia 18:727–733.

Scotet E, Houssaint E. 1998. Exon III splicing switch of fibroblast growth factor (FGF) receptor-2 and -3 can be induced by FGF-1 or FGF-2. Oncogene 17:67–76.

Serra E, Ars E, Ravella A, Sanchez A, Puig S, Rosenbaum T, Estivill X, Lazaro C. 2001. Somatic NF1 mutational spectrum in benign neurofibromas: mRNA splice defects are common among point mutations. Hum Genet 108:416–429.

Shen H, Kan JL, Green MR. 2004. Arginine-serine-rich domains bound at splicing enhancers contact the branchpoint to promote prespliceosome assembly. Mol Cell 13:367–376.

Shibuya Y, Okabayashi T, Oda K, Tanaka N. 1998. Ratio of CD44 epithelial variant to CD44 hematopoietic variant is a useful prognostic indicator in gastric and colorectal carcinoma. Jpn J Clin Oncol 28:609–614.

Stamm S. 2002. Signals and their transduction pathways regulating alternative splicing: a new dimension of the human genome. Hum Mol Genet 11:2409–2416.

Steinman HA, Burstein E, Lengner C, Gosselin J, Pihan G, Duckett CS, Jones SN. 2004. An alternative splice form of Mdm2 induces p53-independent cell growth and tumorigenesis. J Biol Chem 279:4877–4886.

Stickeler E, Kittrell F, Medina D, Berget SM. 1999. Stage-specific changes in SR splicing factors and alternative splicing in mammary tumorigenesis. Oncogene 18:3574–3582.

Sun Z, Pan J, Bubley G, Balk SP. 1997. Frequent abnormalities of TSG101 transcripts in human prostate cancer. Oncogene 15:3121–3125.

ten Dam GB, Zilch CF, Wallace D, Wieringa B, Beverley PC, Poels LG, Screaton GR. 2000. Regulation of alternative splicing of CD45 by antagonistic effects of SR protein splicing factors. J Immunol 164:5287–5295.

Terry G, Ho L, Londesborough P, Cuzick J. 2002. Abnormal FHIT expression profiles in cervical intraepithelial neoplastic (CIN) lesions. Br J Cancer 86:376–381.

Thanaraj TA, Stamm S, Clark F, Riethoven JJ, Le T, V, Muilu J. 2004. ASD: the Alternative Splicing Database. Nucleic Acids Res 32:D64–D69.

Tureci O, Sahin U, Zwick C, Neumann F, Pfreundschuh M. 1999. Exploitation of the antibody repertoire of cancer patients for the identification of human tumor antigens. Hybridoma 18: 23–28.

Usener D, Schadendorf D, Koch J, Dubel S, Eichmuller S. 2003. cTAGE: a cutaneous T cell lymphoma associated antigen family with tumor-specific splicing. J Invest Dermatol 121:198–206.

Van Buskirk C, Schupbach T. 2002. Half pint regulates alternative splice site selection in Drosophila. Dev Cell 2:343–353.

van Doorn R, Dijkman R, Vermeer MH, Starink TM, Willemze R, Tensen CP. 2002. A novel splice variant of the Fas gene in patients with cutaneous T-cell lymphoma. Cancer Res 62:5389–5392.

Wang RF, Appella E, Kawakami Y, Kang X, Rosenberg SA. 1996. Identification of TRP-2 as a human tumor antigen recognized by cytotoxic T lymphocytes. J Exp Med 184:2207–2216.

Wang NM, Chang JG. 1999. Are aberrant transcripts of FHIT, TSG101, and PTEN/MMAC1 oncogenesis related? Int J Mol Med 3:491–495.

Wang NM, Chang JG, Liu TC, Lin SF, Peng CT, Tsai FJ, Tsai CH. 2000. Aberrant transcripts of FHIT, TSG101 and PTEN/MMAC1 genes in normal peripheral mononuclear cells. Int J Oncol 16:75–80.

Wang HY, Xu X, Ding JH, Bermingham JR, Jr., Fu XD. 2001a. SC35 plays a role in T cell development and alternative splicing of CD45. Mol Cell 7:331–342.

Wang T, Fan L, Watanabe Y, McNeill P, Fanger GR, Persing DH, Reed SG. 2001b. L552S, an alternatively spliced isoform of XAGE-1, is over-expressed in lung adenocarcinoma. Oncogene 20:7699–7709.

Wang Z, Lo HS, Yang H, Gere S, Hu Y, Buetow KH, Lee MP. 2003. Computational analysis and experimental validation of tumor-associated alternative RNA splicing in human cancer. Cancer Res 63:655–657.

Wang HC, Su YR, Han KJ, Pang XW, Peng JR, Liang B, Wang S, Chen WF. 2004. Multiple variants and a differential splicing pattern of kinectin in human hepatocellular carcinoma. Biochem Cell Biol 82:321–327.

Xie H, Zhu WY, Wasserman A, Grebinskiy V, Olson A, Mintz L. 2002. Computational analysis of alternative splicing using EST tissue information. Genomics 80:326–330.

Xu Q, Modrek B, Lee C. 2002. Genome-wide detection of tissue-specific alternative splicing in the human transcriptome. Nucleic Acids Res 30:3754–3766.

Xu Q, Lee C. 2003. Discovery of novel splice forms and functional analysis of cancer-specific alternative splicing in human expressed sequences. Nucleic Acids Res 31:5635–5643.

Yan Y, Phan L, Yang F, Talpaz M et al. 2004. A novel mechanism of alternative promoter and splicing regulates the epitope generation of tumor antigen CML66-L. J Immunol 172:651–660.

Yeakley JM, Fan JB, Doucet D, Luo L, Wickham E, Ye Z, Chee MS, Fu XD. 2002. Profiling alternative splicing on fiber-optic arrays. Nat Biotechnol 20:353–358.

Zhou YQ, He C, Chen YQ, Wang D, Wang MH. 2003. Altered expression of the RON receptor tyrosine kinase in primary human colorectal adenocarcinomas: generation of different splicing RON variants and their oncogenic potential. Oncogene 22:186–197.

Zhu J, Mayeda A, Krainer AR. 2001. Exon identity established through differential antagonism between exonic splicing silencer-bound hnRNP A1 and enhancer-bound SR proteins. Mol Cell 8:1351–1361.

Zhu J, Shendure J, Mitra RD, Church GM. 2003. Single molecule profiling of alternative pre-mRNA splicing. Science 301:836–838.

Journal of Immunotoxicology, 2009; 6(4): 227–234

informa
healthcare

RESEARCH ARTICLE

ELISPOT assays provide reproducible results among different laboratories for T-cell immune monitoring—even in hands of ELISPOT-inexperienced investigators

W. Zhang[1], R. Caspell[1], A. Y. Karulin[1], M. Ahmad[2], N. Haicheur[3], A. Abdelsalam[4], K. Johannesen[5], V. Vignard[6], P. Dudzik[7], K. Georgakopoulou[8], A. Mihaylova[9], K. Silina[10], N Aptsiauri[11], V. Adams[12], P. V. Lehmann[1,13], and S. McArdle[2]

[1]Cellular Technology Ltd., Shaker Hts. Ohio, USA, [2]Nottingham Trent University, Nottingham, UK, [3]Hôpital Européen Georges Pompidou, Paris, France,[4]University of Pittsburgh, Pittsburgh, PA, USA, [5]Norwegian Radium Institute, Oslo, Norway, [6]INSERM, Nantes, France, [7]Jagiellonian University Medical College, Krakow, Poland, [8]Immunology Center, St. Savas Cancer Hospital, Athens, Greece, [9]UMBAL "Alexandrovska", Sofia, Bulgaria, [10]Biomedical Research Study Centre, Riga, Latvia, [11]Hospital Universitario Virgen De la Nieves, Granada, Spain, [12]Onyvax Ltd, London, UK, and [13]Department of Pathology, Case Western Reserve University, Cleveland OH, USA

Abstract

Measurements of antibodies in bodily fluids (e.g., by ELISA) have provided robust and reproducible results for decades and such assays have been validated for monitoring of B-cell immunity. In contrast, measuring T-cell immunity has proven to be a challenge due to the need to test live cells in functional assays *ex vivo*. Several previous efforts looking into the reproducibility of *ex vivo* T-cell assays between different laboratories, or even within the same laboratory, have provided rather discouraging results. The hypothesis we tested in this study is that those poor results are due to the lack of assay and data analysis standardization, rather than the inherent complexity of T-cell assays. In this study, 11 laboratories across Europe and the United States were provided identical reagents and were asked to follow the same protocol while testing aliquots of the same three cryopreserved peripheral blood mononuclear cells (PBMC) in an interferon-γ (IFNγ) ELISPOT assay measuring the antigen-specific T-cell response to a CMV peptide. All individuals performing the assays were ELISPOT novices. At their first attempt, while three of these individuals failed with the basic logistics of the trial, eight detected the peptide-specific CD8+ T-cells in frequencies approximating the values established by the Reference Laboratory. The data show that ELISPOT assays provide reproducible results among different laboratories when the assay procedure and data analysis is standardized. Since ELISPOT assays have been qualified and validated for regulated studies, they are ideal candidates for robust and reproducible monitoring of T-cell activity *in vivo*.

Keywords: *Immune monitoring; CD8 cell immunity; immune assay standardization*

Introduction

T-cell immunity is critically involved in combating infections and cancer, as well as playing pathogenic roles in autoimmune disease and allergies. Therefore, monitoring antigen-specific T-cells and their effector functions is critical for the understanding of these diseases and for proper assessments of the efficacies of specific immune therapies. Yet, unlike the detection of antibodies, reliable measurement of T-cell-mediated immunity has, until now, continued to be a major challenge, this resulting from several factors. One such factor is that the antigen-specific T-cells of interest typically occur in very low frequencies in test samples such as peripheral blood. In addition, for an assay to reliably measure T-cell function(s), it needs to be warranted that the test conditions are such that the function of T-cells *in vitro* remains unimpaired relative to *in vivo*. The many variables that can affect T-cell functionality have earned T-cell assays the reputation of being rather fragile, with even minor changes of test conditions potentially having a major impact on the test results.

Address for Correspondence: Paul V. Lehmann, 20521 Chagrin Boulevard, Shaker Heights, OH 44122, USA; E-mail: pvl@immunospot.com.

(Received 27 July 2009; revised 04 September 2009; accepted 08 September 2009)

ISSN 1547-691X print/ISSN 1547-6901 online © 2009 Informa UK Ltd
DOI: 10.3109/15476910903317546

http://www.informahealthcare.com/imt

Further complicating the field of T-cell diagnostics, such assays traditionally have been carried out with freshly-isolated T-cells. While working with freshly-isolated cells may assure unimpaired functionality of the T-cells, it has the distinct disadvantage that assay results cannot be reproduced with the same cell material. Even in cases where the same subject is bled serially, the peripheral T-cell compartment will not be identical. Circadian rhythms and environmental influences such as nutrition, infections, and stress will change the cellular composition of the peripheral blood mononuclear cells (PBMC) and, subsequently, the functionality of the T-cells when tested in different bleeds of the same donor. Thus, when using freshly-isolated cell material, T-cell assay standardization is challenging even within a single laboratory, and assay harmonization among different laboratories has been close to impossible due to the difficulties inherent in transporting live cell material.

A major breakthrough in the field of T-cell monitoring has been the introduction of protocols that facilitate cryopreservation of PBMC such that, upon thawing, the cells retain their full functionality (Kreher et al., 2003). Thus, it has become possible to freeze multiple aliquots of PBMC obtained from a single bleed of a donor, thereby generating "reference PBMC." In this way, one can use identical cell material for repeated testing in the same laboratory, or potentially in several remote facilities. Such cryopreserved PBMC are ideal tools for assay development and standardization, as well as providing the perfect platform for the comparison of T-cell assay performance in different laboratories.

Using such cryopreserved PBMC, it has recently been shown that T-cell assays can be used to precisely and stably quantify antigen-specific T-cell responses in PBMC involving highly trained personnel in tightly regulated laboratories that strictly adhere to standard operating procedures (SOPs) (Maecker et al., 2008). Several attempts have been undertaken to address the question of whether various different facilities are capable of generating reproducible data while testing the same PBMC reference samples (Cox et al., 2005; Janetzki et al., 2008; Britten et al., 2009). All of these efforts showed quite substantial variations in test results supporting the notion that harmonization efforts are needed before T-cell assays become robust enough to generate reproducible data across laboratories, and thus become suitable for immune monitoring in multicenter clinical trials (Britten et al., 2007).

The above-mentioned multicenter studies left open a fundamental question: Are results of functional T-cell assays inherently difficult to reproduce among laboratories due to the complexity of these assays and because of the live cell material being tested, or does the variability result from the different test conditions in the participating laboratories? A common feature of these trials was that expert laboratories were invited to participate and each laboratory was permitted to use their preferred reagents, locally established protocols, and favored data analysis methods. Because many experimental details of T-cell assays have not yet been standardized, these individual laboratories followed substantially variant protocols while performing the assays, possibly explaining the variability in the results.

We tested herein the hypothesis that T-cell assays will provide reproducible and robust data from different laboratories when all pertinent test conditions are standardized. In addition to standardizing the protocol as well as all reagents used in the assays, we also made an attempt to standardize the expertise level of the individuals performing the assays, using a somewhat unconventional approach: the personnel performing the tests for this trial were all selected based on the criterion that they had no previous experience with the T-cell assay in question, this being ELISPOT (although the laboratories in which they work may or may not have such expertise). In this way, we wanted to minimize the chance that habitual variations of trained investigators could influence the results, and to ensure that the protocol provided is indeed closely followed.

The results presented here show that antigen-specific T-cells can be detected with astounding precision by the interferon-γ (IFNγ) ELISPOT assay in different laboratories—even by inexperienced personnel—if the same protocol is followed, using the same reagents, and a user-independent platform for data analysis. Thus, the variability in previously reported studies involving all types of T-cell assays are not inherent consequences of working with live cell material in complex assay systems, but must have resulted from variations in the different approaches within the laboratories.

Materials and methods

Peripheral blood mononuclear cells

The human PBMC used for this study were selected from the ePBMC™ library of Cellular Technology Ltd. (CTL, Cleveland, OH). The PBMC in this library of donors are high-resolution HLA typed, and precharacterized for T-cell reactivity to common recall antigens. The cells had been obtained by leukapheresis, and were frozen using serum-free freezing reagents (CTL-Cryo™, by CTL) in up to 2000 identical aliquots. While the material used for this study has been a donation to facilitate this study, these cells are commercially available, permitting anyone to reproduce the results reported here. To minimize the chance of damage to the PBMC during transit, the cells were shipped to the participating laboratories in liquid nitrogen shippers (MVE Vapor Cryoshipper, Chart Industries, Marietta, GA) which maintain their temperature for more than 2 weeks, and the laboratories were instructed to transfer the cells to liquid nitrogen tanks upon receipt, where they would be stored until tested. The cells were thawed according to the detailed protocol provided by CTL using the serum-free thawing reagent CTL Anti-Aggregate™ Wash Supplement (from CTL), and were counted and re-suspended in the serum-free testing medium, CTL-Test™ (from CTL). Alternatively, as specified in the Results section, the cells were processed and tested with the respective laboratories' serum of choice.

Antigen

The antigen used for this study was the HLA-A-2-restricted cytomegalovirus peptide NLVPMVATV (CMV pp65) (Wills et al., 1996). This peptide was provided as part of the commercially available Reference Sample QC Set from Cellular Technology Ltd.. The antigen was used at the concentration(s) specified in the Results section.

ELISPOT assay

All study participants used the human IFNγ ELISPOT kit by BD Biosciences (San Diego, CA, USA) and followed the manufacturer's instruction sheet for using the kit. The PBMC were tested within 3 hr after thawing (without an overnight resting period). The PBMC were plated at 400,000 cells per well, except for Figure 2 where they were titrated in serial dilution, as specified.

Antigen-induced cytokine production was measured in quadruplicate wells, except for Figure 1 where single wells were tested. The assay results were analyzed on an ImmunoSpot® Series 5 UV Reader (CTL) using its automated software features for user-independent setting of counting parameters (SmartCount™) and gates (AutoGate™).

Statistical analysis

Except for Figure 1, where single wells were tested, for each test condition the test was performed in quadruplicate wells. Means and standard deviations (SD) were calculated for the quadruplicate wells, and are shown in the figures. Figures 5 and 6 show different representations of the same data set—the means and SD for this dataset are shown in Figure 6. Spot counts represent spots counted per well, as counted, without the spot count in the medium control subtracted. When tested serum-free, the spot counts in the medium control were less than 10 spots per well for all figures—raw data for the serum-free medium spot counts are shown in Figure 6. Stimulation indices (SI) were calculated as follows: SI = mean spot number in quadruplicate antigen-containing wells/mean spot number in quadruplicate medium containing negative control wells.

The full-length instructions and protocols provided to the study participants are available as supplementary material accompanying the online version of this article. The participants were asked to run two assays—the first one was intended for practice and the second one was the actual assay. In this paper, we report the results of the first "practice run" because it already produced impressive results—the results of the second assay were essentially the same as of the first one.

Results and discussion

Overall design of trial

In preparation for this trial, we placed particular emphasis on the thorough characterization of reference PBMC distributed to the participating laboratories. The PBMC were high-resolution HLA, typed. From the PBMC library of Cellular Technology Ltd. (CTL), 38 donors bearing the HLA-A2*0201 allele were selected. The PBMC of all 38 donors were tested in an IFNγ ELISPOT assay for reactivity to the A-2-restricted cytomegalovirus peptide NLVPMVATV (CMV pp65) (Wills et al., 1996), at 1 μg/mL. The response level, measured as spot number in ELISPOT, for the different donors ranged from zero to "too numerous to count" (TNTC) (data not shown). From this library, we selected three PBMC for further characterization, one each with a relatively high, medium, and low reactivity level. Antigen dose–response curves were established for these three donors to select a peptide concentration for plateau-level activation, and the frequencies of CMV pp65-reactive cells were defined by serial dilution of the cells in an IFNγ ELISPOT assay (see later). Moreover, we verified through cell separation experiments and intracytoplasmic IFNγ staining that the CMV pp65-induced IFNγ-producing cells are CD8 positive (data not shown).

The cells were repeatedly tested in IFNγ ELISPOT assays in CTL's Reference Laboratory to establish inter- and intra-assay variation within a single laboratory when the test is performed under GLP-compliant conditions under strict adherence to established SOPs. The "reference value" was established as the mean spot number from the different tests performed in the Reference Laboratory.

Cryopreserved PBMC of the three CMV pp65-reactive donors were then shipped to all participating laboratories. The cells of each PBMC donor were from the same blood draw, thereby assuring that identical cell material was tested by all laboratories. The participating laboratories were then asked to defrost and test the cells in an IFNγ ELISPOT assay using the provided protocol and reagents, and the results obtained were compared with reference value as established in the CTL Reference Laboratory.

Antigen dose responsiveness of the reference PBMC

T-cell activation is antigen dose-dependent, whereby the dose–response curve reflects the functional avidity of the T-cell population to the antigen: high-avidity T-cells are stimulated by low concentrations of antigen, low-avidity T cells are stimulated by high concentrations of antigen (Targoni and Lehmann, 1998). In ELISPOT assays, titrating peptide-antigen typically results in a characteristic sigmoidal dose–response curve; the peptide concentration that induces 50% maximal activation representing the K_{eff} value (Hesse et al., 2001). We tested the dose–response curves of three PBMC to CMV pp65 (Figure 1), permitting us to select a peptide concentration for this study that induces maximal stimulation in all three PBMC: this was determined at 0.1 μg/mL.

As this concentration lies along the "plateau" of the does–response curves for all three PBMC, we can thus assure that minor variations in peptide concentration due to operator variability would not affect the results. All subsequent data reported here were obtained using this peptide concentration. The choice of a single peptide antigen, as opposed to the complex Cytomegalovirus, Epstein-Barr virus, and Flu virus (CEF) control peptide pool (Currier et al., 2002) may have contributed to the reproducibility of the data reported

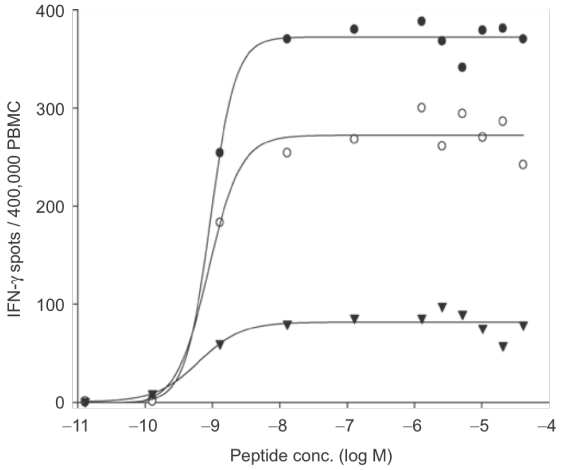


Figure 1. *Dose–response curves of the three reference PBMC after stimulation with CMV pp65.* The cryopreserved cells were thawed and processed under serum free conditions and plated at 4×10^5 cells per well into an IFNγ ELISPOT assay. CMV peptide pp65 was titrated as specified. One well was tested at each peptide concentration. The number of spots induced is shown for each concentration and each sample.

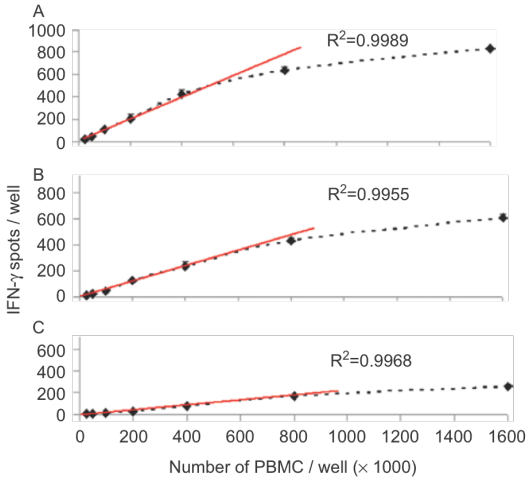


Figure 2. *Relationship between PBMC number plated and IFNγ spots induced by peptide CMV pp65 in the reference PBMC.* The three cryopreserved reference sample PBMC LP-51 (A) LP-37(B) LP-58(C) were thawed and plated in the specified cell numbers per well, in quadruplicate wells. Peptide CMV pp65 was added to each well at 0.1 µg/mL. The mean and standard deviation for each cell number is specified. The regression line was calculated for 4×10^5 to 5×10^4 PBMC per well (red line), with the correlation coefficient shown.

in this study. The individual CEF peptides within the pool trigger T-cells with fundamentally different functional avidities (W. Zhang, unpublished data), and in addition, the individual peptides compete for MHC binding (Lehmann et al., 1989). Thus, while inter-experimental variations in the concentration of the CEF peptide pool used in the preceding multicenter studies might have contributed to the variable T-cell responses reported, in this study the potential stimulatory activity of the antigen was controlled.

Dependence of assay results on cell numbers plated

T-cell activation is dependent on the interaction of a T cell with an antigen-presenting cell. Ideal test conditions therefore will ensure that the two cell types are able to interact. In ELISPOT assays, such contacts are critically dependent on the density of the PBMC in the assay wells. To establish the dependence of ELISPOT assay results on PBMC numbers per well, we plated the reference samples in serial dilutions ranging from 1.6×10^6 to 5.0×10^5 PBMC per well. The peptide was tested at 0.1 µg/mL, and the number of IFNγ spots elicited was determined.

The results are summarized in Figure 2. The number of PBMC plated was found to be in a direct linear relationship to the number of spots detected ($R^2 > 0.995$) in the range between 5×10^4 and 4×10^5 cells per well for all three samples (the regression lines and correlation coefficients calculated for 400,000 to 50,000 PBMC per well are shown in Figure 2). On the basis of this result, a concentration of 4×10^5 cells per well was selected for this study. These data also established that cell counting or pipetting errors affecting the cell numbers plated will also affect the spot counts in direct proportionality. All subsequent data were obtained by testing 4×10^5 cells per well.

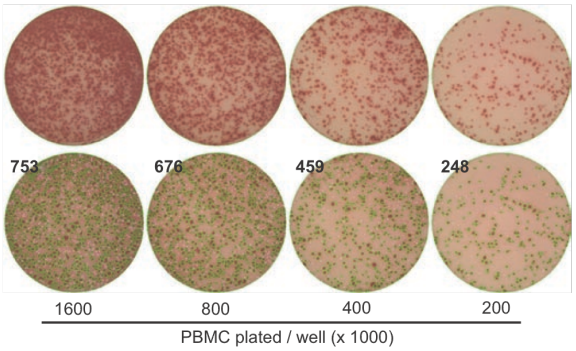


Figure 3. *Upper detection limit of IFNγ spot counting.* Representative images are shown for data obtained with LP-51 when tested at the cell numbers specified. The test conditions are as described in Figure 2. The top row shows the raw images, the bottom row the spot counts with the overlay of the counted spots over the raw image; the green outlines are generated by the ImmunoSpot software to mark spots that have been automatically recognized/counted by the software. The data illustrate how individual spots can be discretely discerned up ≈ 500 spots per well. Note also the linearity of the counts in Figure 2. At higher numbers, confluence of spots and ELISA effects interfere with the countability of the data.

At cell concentrations higher than 8×10^5 per well, the linear relationship between cell and spot number was lost. One reason for this was an observed overlap or confluence of spots in samples where reactivity was high. Figure 3 shows raw as well as counted images for such wells.

As is apparent from the images, discrete spots are observed at frequencies of up to 400 spots per well. Spot confluence

and elevation of background due to "ELISA effects" start to occur at 8×10^5 cells per well in the high responding donor, making spot counts unreliable when they are above 600 spots per well. In this study, spot counts higher than 600 per well were considered "too numerous to count" or TNTC.

Establishing levels of intralaboratory variability for test results and determining a reference value

The PBMC of all three donors were tested in the Reference Laboratory of the CTL (in three different experiments) by a single ELISPOT-trained investigator in a GLP-compliant facility, following established SOPs (Figure 4A). Subsequently, the three PBMC samples were tested by three different ELISPOT- and GLP-trained investigators in the same laboratory following the same SOPs, with each investigator independently processing the samples through all steps of the assay (Figure 4B). In a separate experiment, we attempted to mitigate the effects of variability introduced during the processing of cells and in doing so, assess any potential variability in the performance of the ELISPOT assay itself. For this experiment, the three PBMC were thawed, washed, and counted by a single individual, and then the cells were split for independent testing by three investigators. The results are shown in Figure 4C.

Since each of the five experiments shown in Figures 1, 2, and 4 represent an independent frequency measurement in the Reference Laboratory, we used the results of all five individual experiments to establish the reference values for each of the PBMC as the mean ± SD of the spot counts from the five experiments. The mean ± 3SD for the each of the three reference samples are shown on the left-hand side in Figure 5. Medium controls are not shown, as they were < 10 spots per well in each of these experiments.

Establishing interlaboratory variability of test results

The three cryopreserved PBMC reference samples were distributed to 11 laboratories in 8 countries (e.g., Bulgaria, France, Greece, Latvia, Norway, Poland, Spain, the United Kingdom, and the United States). In addition, all participating laboratories were provided with identical serum-free thawing/washing, and testing media (CTL-Anti-Aggregate-Wash™ and CTL-Test™, respectively), as well as CTL's detailed protocols for thawing and processing the cells. The laboratories received CMV pp65 peptide in a ready-to-use concentration, and were provided a human IFNγ ELISPOT kit (from BD Biosciences). Each participant performed the assays according to the manufacturer's specifications. Importantly, all individuals performing the assay were ELISPOT-inexperienced. The plates were scanned and analyzed centrally at Nottingham Trent University, United Kingdom, using the automated analysis parameters of a CTL ImmunoSpot S5 UV Reader. The results obtained in eight laboratories are shown in Figures 5 and 6. Note that three laboratories failed to provide results: one of these reported that the cells were inadvertently thawed and refrozen before the assay was performed; one switched the washing and testing media during the preparation of the cells for the assay; and the third laboratory reported having mixed up the reference samples.

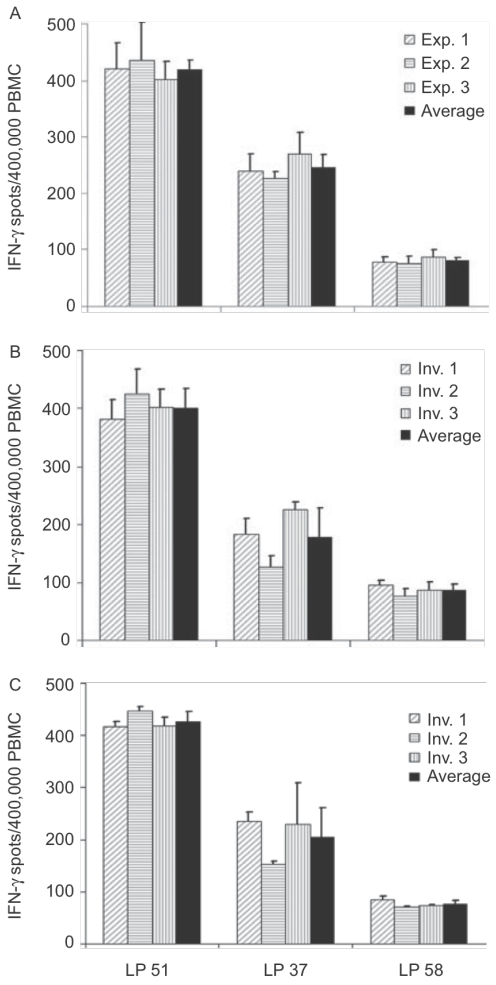


Figure 4. *Intra-laboratory reproducibility of data under GLP conditions.* The three reference samples as specified by shading were tested in three independent experiments by a single individual (A) or by three different individuals, whereby each of them thawed and processed the cells independently (B). In (C), the cells were thawed in a separate experiment and processed by a single individual and subsequently split into three identical aliquots for handling by three independent investigators. The results obtained by each individual are represented by the grey bars (with the SD for triplicate wells) shown; the mean ± SD of the three independently-obtained data points are specified by the black bar. All the data were obtained in the Reference Laboratory while testing under GLP-compliant conditions.

In all eight laboratories, for all three PBMC, the medium background was < 10 spots per well (when tested under serum-free conditions, see Figure 6) and all detected very clear-cut CMV pp65-induced responses (all stimulation indices were >27, shown and discussed later in Figure 6). Most of the spot counts obtained in the different laboratories were within three standard deviations (SD) of the Reference Value provided by the Reference Laboratory

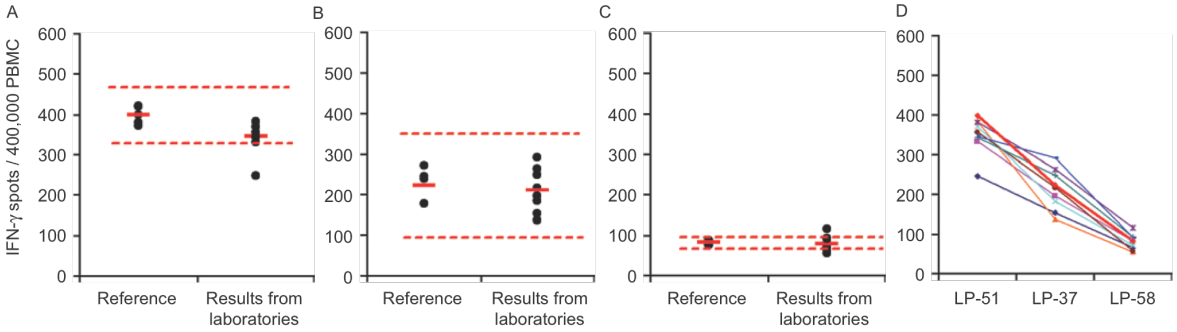


Figure 5. *Inter-laboratory variability of the test results.* Three reference samples shown in the left three panels (A: LP-51, B: LP-37, C: LP-58) were tested independently in eight different laboratories following identical protocols and using identical reagents, including the use of serum free media. Each dot on the right represents the test result obtained by one of these laboratories, with the mean for all laboratories shown by the red bar. The results of the individual tests performed in the Reference Laboratory (from Figures 2, 3, and 4) are represented by the dots on the right, with the mean shown by the red bar (representing the Reference Value for this study), and ± 3 SD shown by the hatched red line. Panel D shows the ranking hierarchy, whereby the results obtained by each laboratory are represented by a different color and are connected by a line in that color.

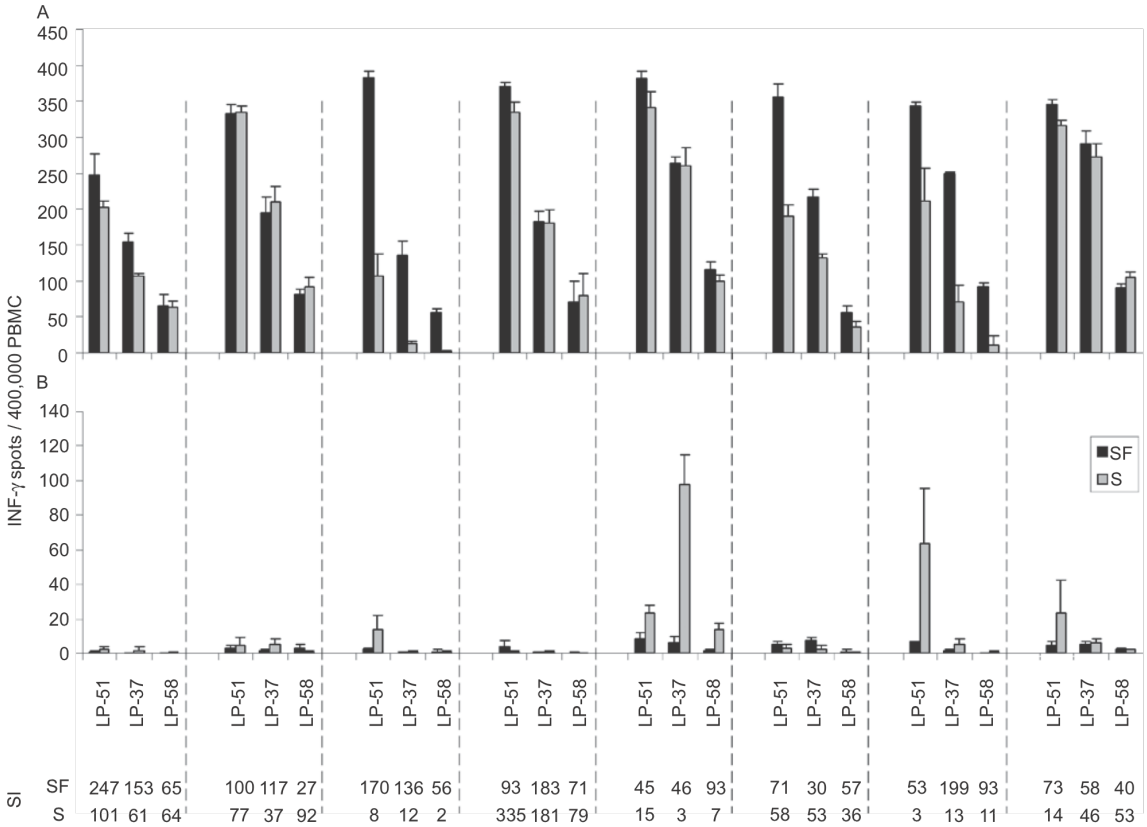


Figure 6. *Assay performance in serum free test conditions versus serum-containing media.* All participating laboratories tested the PBMC reference samples in serum free CTL-Test™ medium (SF, the black bars) or with the serum of their choice (S, the grey bars—these media contained 10% of the respective serum plus 1% L-glutamine in RPMI 1640). Panel (A) shows the CMV pp 65-induced spot counts. Panel (B) shows the spot counts in the respective media alone. In each of the panels, the data for the three reference PBMC (as specified) are grouped per laboratory being separated by the hatched line. Thus, within each hatched line a different serum was tested. Means (±1SD) are shown for each test condition done in quadruplicate wells. The last two rows show the stimulation index (SI) for each test condition serum-free or with the respective serum. SI was calculated as spot counts induced by antigen divided by spot count in the medium control. The data in this graph for the SF media are the same as in Figure 5.

(Figures 5A–5C). Moreover, each of the laboratories correctly established the response hierarchy/rank among the PBMC samples, with spot counts for LP-51 > LP-37 > LP-58 (Figure 5D).

Furthermore, deviations from the reference value tended to be systematic, that is, if one laboratory detected, for example, a lower spot number for one PBMC sample relative to another laboratory, the values for the other two PBMC tended to be also lower (Figure 5D). A likely explanation for such systematic deviations is to be sought in cell counting. Therefore, it would seem that, with more standardization for cell counting, and some practice with the technique, these already impressive data on interlaboratory reproducibility of results will move yet closer to agreement with each other. It should be reiterated that the individuals performing the assays were ELISPOT-inexperienced, and that these results represent their very first attempt at performing the assay.

Previous attempts to show reproducibility of T-cell assays using ELISPOT, Tetramers/Pentamers, ICS, and CFSE dilution indicated that there were substantial inter-laboratory variations (Cox et al., 2005; Janetzki et al., 2008; Britten et al., 2009). Because in those attempts each participating laboratory was permitted to use its favorite reagents and test conditions, it remained unclear whether T-cell assays are inherently hard to reproduce, or whether the variable outcomes resulted from the different test conditions used in the participating laboratories. The study here clearly answers this question for ELISPOT, and invites similar comparisons for other T-cell assays.

Influence of serum on the test results

Cellular assays in general, and T-cell assays in particular, have been traditionally performed using serum-containing media that rely on serum to provide the essential microenvironment for maintaining the viability and functionality of PBMC in culture. Only recently have serum-free media become available that are suited for cryopreservation and *ex vivo* testing of PBMC, such as the medium used to generate the above data. In addition to testing the samples on this common platform, with the same serum-free reagents used for thawing, washing (CTL Anti-Aggregate Wash™), and testing of the PBMC (CTL-Test™), the participating laboratories were asked to do a parallel experiment in which the cells were processed and tested using the serum of their choice—a serum which had been used in their respective laboratories for cell culture/T-cell assays. Thus, since in the same experiment, under otherwise identical conditions, the cells were tested with eight different sera vs. in the same serum-free medium, one can directly compare the assay performance under these conditions. As shown in Figure 6A, the spot counts in CTL-Test™ (the black bars) were either similar (Laboratories 1, 2, 4, 5, and 7) or significantly higher (Laboratories 3, 6, and 7) than those induced in the different serum-containing media. The medium background tended to be higher in the serum-containing media, and was disturbingly high in the serum of Laboratories 5 and 7 (Figure 6B). This resulted for most laboratories in an increased signal-to-noise ratio of assay performance in favor of the serum-free medium, expressed as stimulation index (SI = spot counts induced by antigen/spot counts in the medium control; see last two rows in Figure 6). These data establish that—while reproducible results among laboratories can be obtained using serum-free media—different sera used in the different laboratories can introduce variability in the assay results. The fact that serum can be a major assay variable might explain why this study, performed using serum-free conditions, showed much higher reproducibility of test results among laboratories than was seen in the previous attempts, in which each laboratory was permitted to use its own favored serum.

The higher reproducibility of results in serum-free media could benefit clinical trials in which there are multiple testing sites. For example, while the spot counts for LP-37 in serum-free media ranged from 130 to 300 spots, due to the low medium background, all laboratories reported strongly positive SI (ranging from 30 to 199; see Figure 6). In contrast, for the serum-containing media one of the laboratories recorded for LP-37 a borderline (SI = 3) and two laboratories a weakly-positive result with SI =12 and 13, respectively (Figure 6).

Concluding remarks

For the scientific community, the notion that T-cell assays can be reliably performed, only in expert laboratories by highly-trained personnel, has become accepted wisdom. This notion, however, also implies that T-cell assays are inherently complex and their results hard to reproduce, which caused a large segment of the non-T-cell specialized research community to refrain from utilizing T-cell monitoring techniques. Pharmaceutical industry and governmental agencies have generally been reluctant to include T-cell assays in their immune monitoring efforts as well because it was believed to be very difficult—if not impossible—to validate and qualify T-cell assays. Regulatory agencies will only accept qualified assays for use in development of new drugs, or as surrogate clinical end points (van der Burg, 2008). Recently, ELISPOT assays have been qualified and validated in laboratories adhering to GLP, including in the reference laboratory of this study. The data reported here, showing reproducibility of T-cell assay results between different laboratories, may boost the confidence that, in the near future, T-cell assays will join antibody measurements for robust assessments of immunity both by regulatory agencies and a broader, non-T-cell-specialized scientific community.

Acknowledgments

We thank Cellular Technology Ltd. (CTL) for donating their reference PBMC, CMV pp65 peptide, and the serum-free wash- (CTL Anti-Aggregate-Wash) and test-media (CTL-Test). We also thank BD Biosciences for donating their human IFNγ ELISPOT kits. Most of all, we thank the directors of the laboratories who supported this project by assigning an ELISPOT inexperienced member of their staff to this

project, namely Drs. C. Baxevanis, A. DeLeo, G. Gaudernack, F. Garrido, F. Jotereau, P. Laidler, A. Line, E. Tartour, E. Naumova, and R. Rees.

***Declaration of interest*:** Dr. Lehmann is employed by CTL. CTL is the manufacturer of ELISPOT readers, of the PBMC, and serum-free media used in this study. Independent evaluation of the study was conducted by Dr. McArdle. The Authors alone are responsible for the content and writing of this paper. The Authors did not receive funding for this study.

References

Britten, C. M., Janetzki, S., van der Burg, S. H., Gouttefangeas, C., and Hoos, A. 2007 Toward the harmonization of immune monitoring in clinical trials: *Quo vadis*? *Cancer Immunol. Immunother*. 57:285–288.

Britten, C. M., Janetzki, S., Ben-Porat, L., Clay, T. M., Kalos, M., Maecker, H., Odunsi, K., Pride, M., Old, L., Hoos, A., and Romero, P. 2009. Harmonization guidelines for HLA-peptide multimer assays derived from results of a large-scale international proficiency panel of the Cancer Vaccine Consortium. *Cancer Immunol. Immunother*. DOI 10.1007/s00262-009-0681-z.

Cox, J. H., Ferrari, G., Kalams, S. A., Lopaczynski, W., Oden, N., and D'souza, M. P. 2005. Results of an ELISPOT proficiency panel conducted in 11 laboratories participating in international human immunodeficiency virus type 1 vaccine trials. *AIDS Res. Hum. Retrovirus*. 21:68–81.

Currier, J. R., Kuta, E. G., Turk, E., Earhart, L. B., Loomis-Price, L., Janetzki, S., Ferrari, G., Birx, D. L., and Cox, J. H. 2002. A panel of MHC Class I-restricted viral peptides for use as a quality control for vaccine trial ELISPOT assays. *J. Immunol. Meth*. 260:157–172.

Hesse, M. D., Karulin, A. Y., Boehm, B. O., Lehmann, P. V., and Tary-Lehmann, M. 2001. A T-cell clone's avidity is a function of its activation state. *J. Immunol*. 167:1353–1361.

Janetzki, S., Panageas, K. S., Ben-Porat, L., Boyer, J., Britten, C. M., Clay, T. M., Kalos, M., Maecker, H. T., Romero, P., Yuan, J., Kast, W. M., and Hoos, A. 2008. Results and harmonization guidelines from two large-scale international ELISPOT proficiency panels conducted by the Cancer Vaccine Consortium (CVC/SVI). *Cancer Immunol. Immunother*. 57:303–315.

Kreher, C. R., Dittrich, M. T., Guerkov, R., Boehm, B. O., and Tary-Lehmann, M. 2003. $CD4^+$ and $CD8^+$ cells in cryopreserved human PBMC maintain full functionality in cytokine ELISPOT assays. *J. Immunol. Meth*. 278:79–93.

Lehmann, P. V., Cardinaux, F., Appella, E., Muller, S., Falcioni, F., Adorini, L., and Nagy, Z. A. 1989. Inhibition of T-cell response with peptides is influenced by both peptide-binding specificity of major histocompatibility complex molecules and susceptibility of T-cells to blocking. *Eur. J. Immunol*. 19:1071–1077.

Maecker, H. T., Hassler, J., Payne, J. K., Summers, A., Comatas, K., Ghanayem, M., Morse, M. A., Clay, T. M., Lyerly, H. K., Bhatia, S., Ghanekar, S. A., Maino, V. C., Delarosa, C., and Disis, M. L. 2008. Precision and linearity targets for validation of an IFNγ ELISPOT, cytokine flow cytometry, and tetramer assay using CMV peptides. *BMC Immunol*. 9:9.

Targoni, O. S., and Lehmann, P. V. 1998. Endogenous myelin basic protein inactivates the high avidity T-cell repertoire. *J. Exp. Med*. 187:2055–2063.

van der Burg, S. H. 2008. Therapeutic vaccines in cancer: Moving from immunomonitoring to immunoguiding. *Expert Rev. Vaccines* 7:1–5.

Wills, M., Carmichael, A., Mynard, K., Jin, X., Weekes, M., Plachter, B., and Sissons, J. 1996. The human cytotoxic T-lymphocyte (CTL) response to cytomegalovirus is dominated by structural protein pp65: Frequency, specificity, and T-cell receptor usage of pp65-specific CTL. *J. Virol*. 70:7569–7579.

3.6 The analysis of antigen-specific spontaneous CTL responses

Spontaneous T cell immune response against an antigen indicates to an increased possibility to induce a response against that antigen also in therapy. The molecular and immunological techniques used to analyse spontaneous as well as therapy-induced T cell responses are mostly based on the assessment of T cell proliferation potential, produced cytokine profile and the ability to lyse the antigen positive target cells. The analysis of antibody responses has been routinely used by infectologists already for a long time with the most popular method being ELISA, and one of the most widely used T cell activation assays is based on the same principle and is called ELISPOT. This assay can give information also about the functional and cytolytic activity of tested T cells by utilising various cytokine as well as granzyme B and/or perforin-specific detection antibodies[127].

We collected PBMCs from five healthy donors, 11 breast and 11 gastric cancer patients. A pool of viral peptides and the well-known immunogenic CT antigen NY-ESO-1 were chosen to develop T cell activation protocol, and two of the promising candidate antigens LRRC50 and SPAG8 were selected for spontaneous CTL response analysis by using IFNγ ELISPOT (Mabtech, Sweden) in patients with upregulated antigen expression in their tumour tissues or with autoantibody response.

3.6.1 The optimisation of T cell pre-sensitisation assay

As tumour antigen-recognising T cells tend to be very rare in patients blood, *ex vivo* (directly from blood) ELISPOT has been reported to work poorly, thus T cell pre-sensitisation with antigenic peptides is suggested to amplify the specific effectors. However, several different T cell pre-sensitisation protocols exist. Major variations are introduced by (i) the cytokine usage (IL-2 alone or in combination with IL-7), (ii) number of peptide restimulation cycles (one versus several (2 and more), with or without additional APCs), (iii) T cell numbers per stimulation (from $5x10^5$ total PBMCs to $5x10^5$ of purified CD8+ T cells), (iv) peptide purity, length and final concentration in the stimulation, as well as (v) target cell types in ELISPOT (EBV transformed B cells, PHA stimulated CD4+ cells, autologous PBMCs, T2 cell lines etc.), not to mention (vi) reagents from different manufacturers that can influence the outcome of the stimulation. We took part as one of the international partner laboratories in a project aimed at the development of standardized ELISPOT assay and the obtained results are reported in the original paper IV. However this project didn't address the variations in the pre-sensitization step of the T cell analysis.

The initial protocol was designed to compare the impact of the number of peptide stimulations on the expansion of antigen specific effector CTLs in PBMCs of five healthy donors. Three stimulations were compared to two and to one and the results were determined by IFNγ ELISPOT using pool of peptides from CMV, EBV and flue viruses (CEF) (CTL Ltd, USA). It is evident that a single stimulation yields the highest number of reactive cells comparing to two and three stimulations with viral control peptides, and was chosen as the further work setting (Figure 8A). Secondly, it demonstrated that the stimulated cells have expanded significantly comparing to the non-stimulated samples (CD8+ T cells extracted from freshly thawed PBMCs on the day of ELISPOT) showing that the chosen pre-sensitisation setting is appropriate for effector CTL *in vitro* propagation (Figure 8A).

However, the number of viral peptide-specific CTLs in PBMCs is relatively very high (5-20%) and this analysis doesn't grant the detection of rare cancer antigen-specific CTLs (<0.01%). Hence we used one of the most immunogenic tumour antigens NY-ESO-1 to test this protocol on a NY-ESO-1 serum-positive gastric cancer patient and five healthy donors for the presence of specific CTLs. Figure 8B displays the obtained results from

duplicated IFNγ ELISPOT wells from two separate experiments. The CTL activation was observed in one healthy donor, indicating that the elaborated experimental setting allows to detect the presence of rare tumour-antigen specific CTLs, however it is not clear if effector/memory or naïve CD8+ T cells were activated.

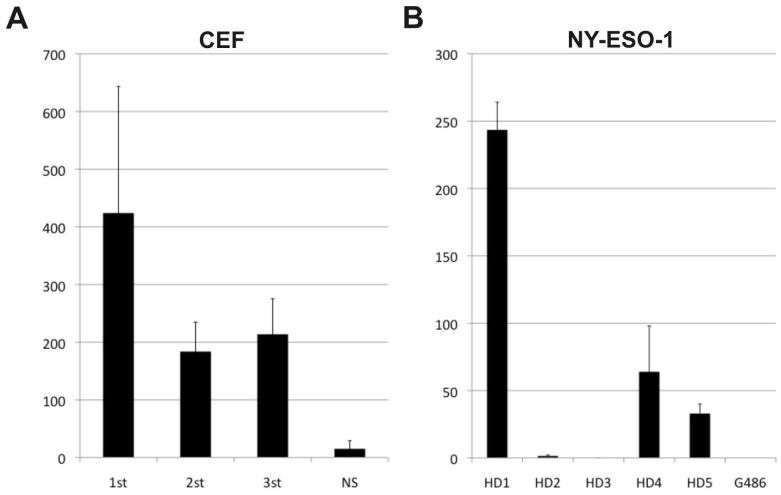


Figure 8. Optimisation of T cell activation protocol.

CD8+ T cells were separated from total PBMCs using magnetic beads and used for peptide stimulations. T cell activation was analysed after 10 days of culture by IFNγ ELISPOT, which was developed after overnight incubation using biotinilated anti-IFNγ detection antibody and Streptavidin-SAP conjugate with BCIP/NBT-plus substrate for colour development. Plates were scanned and analysed by the CTL Europe ScAnalysis service using ImmunoSpot 5.0.3. software. The spot number in the negative control wells was subtracted from the spot number in the stimulation wells. Threshold of CTL activation was set as twice the number of spots in the negative control wells plus ten spots. Y axis indicates spot numbers. **A** $4x10^4$ CD8+ T cells from five healthy donors were seeded in each of three wells for each donor and stimulated with irradiated autologous CD8-CD4- PBMCs ($4x10^5$) pulsed with positive control viral peptide pool (CEF) (1μM each peptide) on culture day 1. Second and third wells were stimulated with peptide pulsed APCs additionally for either one or two times on days 4 and 7, respectively. After 10 days 1/10th of the cultured cells were seeded in each of three IFNγ ELISPOT plate wells and peptide pulsed autologous PHA-stimulated CD4+ T cells (10^5/well) were added as target cells in serum free medium to two wells for T cell activation analysis, and target cells without peptide were added to the third well for negative control. Colons represent the average spot number of the five donors and error bars indicate standard error of the mean among the five donors. X axis represents the number of stimulations per culture: 1st – single stimulation on culture day 1, 2st – two stimulations, with the second stimulation on culture day 4, 3 – three stimulations, with the third stimulation on culture day 7, NS – non-stimulated freshly thawed and separated CD8+ T cells. **B** $4x10^5$ CD8+ T cells from five healthy donors and one gastric cancer patient were stimulated with irradiated autologous CD8-CD4- PBMCs ($4x10^5$) pulsed with NY-ESO-1 peptide pool (1μM each peptide) and cultured for 10 days. After 10 days 1/3 of the cultured cells were seeded in each of three IFNγ ELISPOT plate wells and peptide pulsed autologous PHA-stimulated CD4+ T cells (10^5/well) were added as target cells in serum free medium to two wells for T cell activation analysis, and target cells without peptide were added to the third well for negative control. Colons represent the average spot number from two experiments for each individual. Error bars indicate standard deviation of the replicates. X axis represents patient and healthy donor codes.

3.6.2 Selection of patients for T cell activation analysis

To select patients for the analysis of spontaneous CTL responses we looked for those showing (i) overexpression of LRRC50, SPAG8 or NY-ESO-1 mRNA in tumours by using qPCR (Figure 9), and (ii) the presence of corresponding autoantibodies in their sera by using phage-displayed antigen microarrays (Table 6). A gene was considered as overexpressed in a tumor sample if the expression value in the tumor sample exceeded that of the adjacent normal sample by at least two fold and the average value of all normal samples by at least three fold to account for the interindividual gene expression variations. A prominent overexpression of NY-ESO1 was observed in two breast cancer patients (Figure 9A) of stages I and IIIA (Table 3). We didn't detect any expression of NY-ESO-1 in any available gastric cancer specimens (data not shown). The upregulation of SPAG8 was observed in two breast cancer specimens with early stage disease (I and II), while no prominent overexpression was seen in gastric cancer tissues for which the paired PBMC samples were available (Figure 9B), however the patient with the highest SPAG8 level in the tumour tissue was selected for T cell activation assay. LRRC50 was elevated in cancer

tissues of four patients with breast cancer and three patients with gastric cancer (Figure 9C) stages I-IV. Table 6 displays the summary of obtained mRNA expression and autoantibody data, and the patients selected for T cell activation analyses of the corresponding antigens.

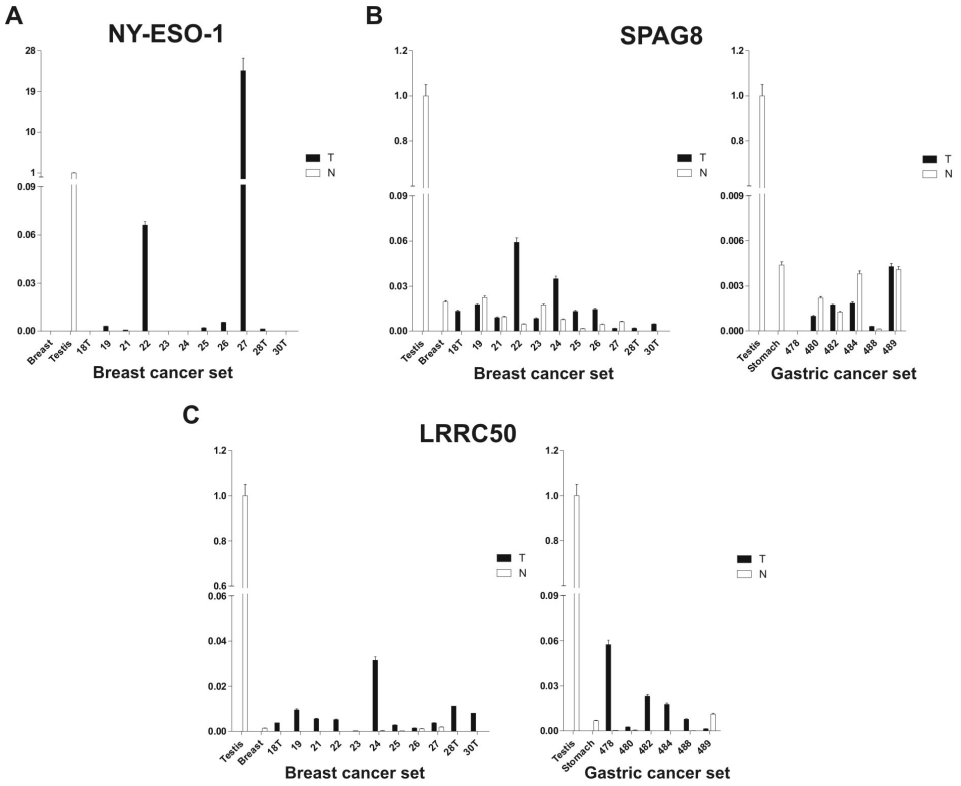


Figure 9. The mRNA expression of cancer antigens in tumour-normal sample pairs of patients with collected PBMCs samples.

QPCR was performed on cDNA samples obtained from tumour and adjacent normal tissues of gastric and breast cancer patients for NY-ESO-1 (A), SPAG8 (B) and LRRC50 (C). The average expression values obtained from two replicates are shown relative to the amount observed in the testis sample; error bars designate standard deviation between the replicates. Samples referred as Testis, Breast and Stomach represent commercial normal samples.

Table 5. Summary of mRNA expression and autoantibody data in patients with collected PBMC samples.

Cancer type	Patient number	Overexpression in tumour			Serum autoantibodies			T cell activation assay
		NY-ESO1	SPAG8	LRRC50	NY-ESO1	SPAG8	LRRC50	
Breast	18	-	-	-	-	-	-	-
	19	-	-	+	-	-	+	LRRC50
	21	-	-	-	-	-	-	-
	22	+	+	-	-	-	-	NY-ESO-1, SPAG8
	23	-	-	-	-	-	-	-
	24	-	+	+	-	-	-	SPAG8, LRRC50
	25	-	-	-	-	-	-	-
	26	-	-	-	-	-	-	-
	27	+	-	-	-	-	-	NY-ESO-1
	28	-	-	+	-	-	+	LRRC50
	30	-	-	+	-	-	-	LRRC50
Gastric	478	-	-	+	-	-	-	LRRC50
	479	NA	NA	NA	-	-	-	-
	480	-	-	-	-	-	-	-
	482	-	-	+	-	-	-	LRRC50
	483	NA	NA	NA	-	-	-	-
	484	-	-	+	-	-	-	LRRC50
	485	NA	NA	NA	+	-	-	NY-ESO-1
	486	NA	NA	NA	+	-	-	NY-ESO-1
	487	NA	NA	NA	-	-	-	-
	488	-	-	-	-	-	-	-
	489	-	-	-	-	-	-	SPAG8

3.6.3 Analysis of naturally occurring antigen-specific CD8+ T cells in breast and gastric cancer patients.

We selected the patients indicated in Table 4 as well as five healthy donors for the LRRC50, SPAG8 and the control antigen NY-ESO-1-specific naturally occurring CTL activation analyses by using IFNγ ELISPOT using the CEF peptide pool as the positive control. We didn't detect any plausible LRRC50 or NY-ESO-1-specific CTL responses, but this initial analysis showed SPAG8-specific CD8+ T cell amplification in one gastric cancer patient (G489) (Figure 10). None of the healthy donors showed any response to the tested cancer antigens (data not shown) except for NY-ESO-1 (Figure 8B).

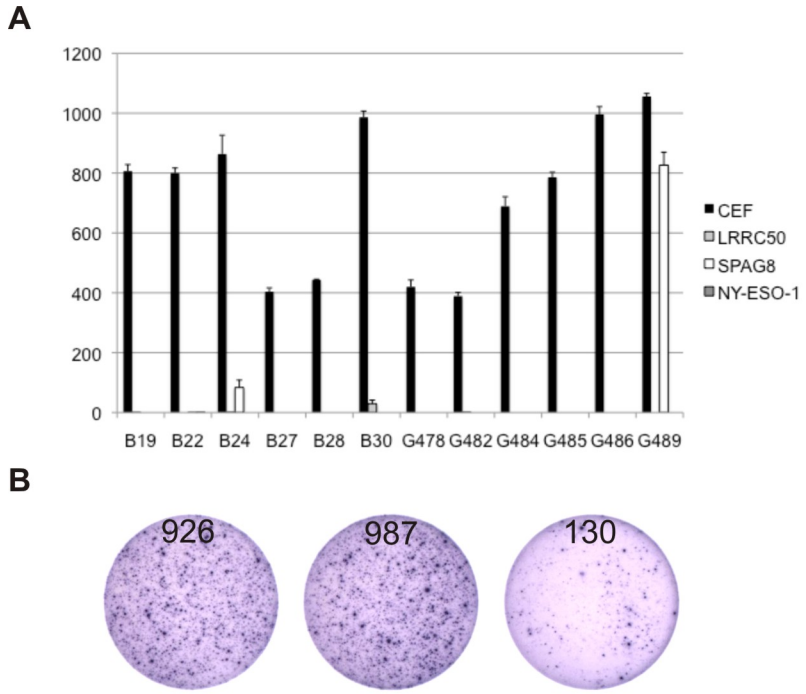


Figure 10. The analyses of LRRC50, SPAG8 and NY-ESO1-specific CD8+ T cells in mRNA or antibody-positive cancer patients.

Purified CD8+ T cells ($4x10^5$) were stimulated with irradiated autologous CD8-CD4- PBMCs ($4x10^5$) pulsed with LRRC50 and NY-ESO-1 peptide pool (1μM each peptide) and cultured for 10 days. After 10 days the cultured cells were split in three parts and seeded in each of three wells of the IFNγ ELISPOT plate. Peptide pulsed autologous PHA-stimulated CD4+ T cells (10^5/well) were added as target cells in serum free medium to two wells, and target cells without peptide were added to the third well for negative control. ELISPOT was developed after overnight incubation using biotinilated anti-IFNγ detection antibody and Streptavidin-SAP conjugate with BCIP/NBT-plus substrate for colour development. Plates were scanned and analysed by the CTL Europe ScAnalysis service using ImmunoSpot 5.0.3. software. The spot number in the negative control wells was subtracted from the spot number in the stimulation wells. Threshold of CTL activation was set as twice the number of spots in the negative control wells plus ten spots. **A** Colons indicate the average values obtained from two replicates, error bars designate standard deviation between the replicates. Y axis indicates spot numbers. X axis shows patient codes; B – breast cancer, G – gastric cancer. **B** The stimulated T cells (left and centre wells) and negative control (right well) of patient G489 cropped out from the scanned image of the developed IFNγ ELISPOT plate with counted spot numbers.

4 Discussion

During the adaptive immune response effector lymphocytes engage in a complex mutual regulatory network, which ensures the fine-tuning of the dominant effector mode differentiating mostly into type 1 and type 2 responses, as well as regulates the extent and duration of the immune response. Such regulatory loops are necessary to fine-tune the defence against the vast universe of pathogens and to ensure the best balance between the destructive and self-preserving capacities of the immune system[128], sort of a "most benefit, least damage" principle. Recent studies have revealed a cornucopia of inter-relationships between adoptive lymphocytes and tumour cells, tumour stroma and cells of the innate immunity that are important for the course of malignancy resulting in either predominantly preventive or promoting influences as dictated by the polarisation into type 1 or type 2 responses, respectively[103]. For example, Th1 cells have evolved to fight intracellular pathogens by promoting CTL activation at the site of infection, while Th2 cells are the major regulators of extracellular pathogen eradication promoting IgE class antibody production and attracting eosinophils and granulocytes – potent inflammatory effectors of the innate immunity to the infection site, hence the opposing contribution to cancer development by determining cytolytic versus inflammatory environments[103]. Thereto the type 1 or Tc1 (conventional) CTLs have been recognised as more potent effectors of tumour cell eradication than the Tc2 cytotoxic CD8+ T cells[129] as they showed prolonged survival, efficient cytotoxicity and skewing Th cells towards type 1 response[130]. However the type 2 Th and CTL effectors have also been shown to be important for tumour clearance in certain circumstances[76]. These data collectively indicate that one needs not only to induce an immune response against a tumour, but also trigger the "right type" of immune response in the right context to achieve a clinical benefit[106].

CTLs have been in the spotlight in attempts to induce a protective cytolytic response owing to their perforin and granzyme B mediated tumour cell killing abilities important in fight against all tumours, but especially those that have downregulated cell-cell contact-dependant death receptor (as FAS, TRAIL, etc.) signalling. A lot of studies have concentrated to the identification of antigens and their CTL epitopes for immune stimulation. However, CTLs themselves don't posses a considerable capacity to induce further amplification of immune response due to various intrinsic self-tolerance mechanisms as AINR[27] etc. This together with the local suppressive tumour microenvironment are mostly the reasons why initial immunotherapies utilising simply MHCI-binding peptides as CTL epitopes or the administration of antigen-specific *in vitro* propagated CD8+ T cells didn't deliver the expected hopes.

The understanding of Th cells as the orchestrators of both, the differentiation and amplification of the immune response, has shifted the attention towards CD4+ T cell adaptive therapy that has a potential to provide polyclonal, integrated B cell and CTL activation[131]. Besides, a direct cytolytic activity of CD4+ T cells has been readily demonstrated, hence these cells could provide a substantial improvement in current immunotherapy statistics[132]. The scene has become increasingly complicated, however, with the identification and characterisation of a myriad of functionally distinct Th cell subsets as well as the mechanisms of DC maturation into various phenotypes that can differentially influence effector cells[133]. For example, activated Th17 cells that have been considered as important players in organ-specific autoimmune diseases are the major producers of IL-21, a cytokine shown to have positive effect on anti-tumour CTLs[134], and the differentiation of this subset is dependant on high TGFβ levels that is a major suppressive cytokine[135]. Another recent capacity of type 2 cytokine IL-4 together with TGFβ is to inhibit the differentiation of Treg cells that are important anti-tumour immunity

suppressors, but instead induce the formation of IL-9 producing effector cells or Th9 cells[136] that again, depending on the context, can be regulatory or inflammatory[133].

Also B cells have provided new surprises starting from the discovery of regulatory subtypes to their importance in Th cell stimulation besides their primary task as antibody producers[68,137], and no longer should be referred to as "B for boring". Analogously to T cells, they can modulate dendritic cells by promoting or inhibiting their maturation as well as skew Th cell subset differentiation at the tumour site depending on the secreted cytokines and antibodies, and again the tumour promoting capacity is more ascribed to the Be2 subtype cells, while protective – to the Be1 type effectors[76,138]. The role of anti-tumour antibodies, however, is under intense debate. These antibodies have been correlated with pro-tumour and pro-inflammatory roles, poor prognosis and increased tumour burden[76], as well as to quite contrary observations related to increased survival and better prognosis[139]. It should be bared in mind that the outcome of antibody presence in the tumour microenvironment is dictated by (i) the secreted Ig class subtype and the balance of activating versus inhibitory Fc receptors on other effector cells recognising antibody-containing immune complexes on tumour cells[70] and of course by (ii) the surrounding cytokine milieu[103]. Here mentioned players are just a tip of the iceberg of the complex nature of the immune response involving also the innate mechanisms, and hence also the concept of the "right adjuvant", or closer the "right cocktail of adjuvants" to activate DCs and mount the protective type immune response together with dampening of suppressor cell activities is crucial for successful immunotherapy development[106].

Besides the "right immune response" and the "right adjuvants" the importance also lays on the target itself – "the right antigen" or the combination of those must be recognised in order to ensure tumour eradication without inducing autoimmunity and avoiding tumour escape[106]. The comprehension of the destructive power of the adaptive immunity as well as the initial animal experiments with chemically inducible tumours suggested that the most suitable anti-tumour targets are the mutated neoantigens formed in cancer due to profound genetic instability because (i) these are tumour-specific without the risk of destroying healthy tissues and (ii) they wouldn't be centrally tolerated[140]. A shift in this paradigm came in the middle of 90'ties when melanoma rejection in patients was shown to be mediated by wild-type melanocyte protein-specific CD8+ T cells, as well as studies demonstrating the manifestation of vitiligo only for patients with regressing tumour, providing evidence for the induction of protective immune response against the so-called tumour-associated antigens that are self-proteins overrepresented in tumours[141]. Tumour-associated antigens possess a threat to induce autoimmunity, hence only antigens shared between tumours and immunoprivileged sites or non-vital organs might be considered for therapy targeting by forced blocking of tolerance and bearable collateral damage[142]. The major advantage of the tumour-associated over tumour-specific antigens is that they are shared among many tumours and may provide the opportunity to develop a cost-effective cancer vaccine, while tumour-specific antigen targeting implies (i) the necessity to characterise each patient's tumour as well as to generate the tumour-specific targets that most often are unique to each patient [1] or (ii) the application of whole tumour vaccines that don't allow the monitoring of therapy induced-immune responses. However, until now there is no consensus on which of these two antigen categories are more suitable for the development of immunotherapy used in every-day practice. Recently guidelines for the prioritisation of antigens applicable in clinical trials have been published and state that the ideal therapy target is both, tumour-specific and shared among many tumours[6] emphasising the need to balance the beneficial treatment outcome with cost-effectiveness and labour-intensiveness. The top five antigens rated using these guidelines developed by a

large panel of experts are the CT antigen WT1, followed by MUC1, two viral antigens LMP2 and HPV E6 E7, and the tumour-specific antigen EGFRvIII[6].

In this work we have tried to approach the "right antigen" problem. Knowing that in order to produce IgG class-switched antibodies a B cell has to receive proper help from a CD4+ T cell that has been activated by the same antigen, and considering the involvement of B cells in modulation of T cell activities, the antigens identified by the utilisation of IgG antibody-based approaches may yield important therapeutic targets. Studies utilising cDNA expression library screening with cancer patients' sera or the SEREX approach have yielded the identification of therapeutically significant cancer antigens like NY-ESO-1[143], and provide ground for the usefulness and potential of this system to discover novel immunotherapy candidate genes. As well the fact that a spontaneous B cell response is observed against antigens currently used in immunotherapy trials indicates that antigens eliciting humoral response might have a better chance to induce immune response also in therapy. We applied phage-display based SEREX to identify a comprehensive set of tumour antigens that could be used in early diagnosis, prognosis and prediction of therapy outcome for melanoma, breast, gastric, and prostate cancers ([144] and unpublished results), which resulted in the identification of 1328 antigenic clones representing 1158 non-redundant antigens. In this study we attempted to characterise the identified antigens and determine whether there might be any therapeutic targets among them that correspond to the established criteria for a potential immunotherapy target antigen[6].

The antigen collection is comprised of conventional natural ORF antigens (15%), undefined ORF antigens (2%) that include uncharacterised splice variants and novel genes, and clones translated as unnatural peptides (83%). Such a large number of unnatural ORF antigens has also been reported by other groups using phage-display technology[145] and to a lesser extent also conventional SEREX. We looked into more detail of the recognition of these peptides to see if the corresponding antibodies indeed react with this unnatural product of a gene and not with its natural ORF counterpart that could had been displayed on the phage due to possible mutations in the phage genome causing translational shift, and we showed that the unnatural peptides were undeniably recognised by the reacting sera (see original paper I).

Around one half of the unnatural peptides are derived from inserts of non-cDNA nature including intergenic regions despite stringent DNAse treatment prior to cDNA library construction[144], which might indicate to a broader transcriptome than previously anticipated supporting the recent acknowledgement of the wide world of non-coding RNAs[146,147]. It is not known whether any of the non-coding, particularly the long non-coding, RNA species could actually be used by the translation machinery as a result of deregulation in cancer and most likely represent mimotopes of other antigens like proteins with changed structure including altered PTMs or lipid and carbohydrate epitopes; the latter two represent mostly B-cell epitopes[148].

The concept of mimotopes developed subsequently after the realisation that antibody cross-reactivity is a common phenomenon because a limited degree of similarity between an epitope and a paratope is sufficient to allow the peptide to bind to an anti-protein antibody[149]. The great proportion of mimotopes among the identified phage-displayed antigens might be explained by the fact that the vast majority of protein epitopes are the so-called conformational or discontinuous epitopes dictated by the natural conformation of the protein[149], and there are several reasons why the probability for an antibody to meet its mimotope in this system is much higher than to meet its true discontinuous epitope: (i) it is often that these epitopes are not preserved on the recombinant phage because usually only a fragment of the whole protein is cloned as well as due to the lack of eukaryotic PTMs in the bacterial cloning system, and (ii) there is a plethora of out-of-frame peptides in the

phage-displayed cDNA library due to the randomness of cloning creating a sufficient diversity of peptides resembling the true discontinuous and PTM-mediated epitopes. This also answers the question why there are so few natural ORF antigens among the identified clones – as there is only a fragment of a protein displayed on the recombinant phage this antigen might more often represent true linear epitopes, which are per se more rare than the discontinuous ones. This notion is also supported by our results showing that only one out of eight tested phage-displayed natural ORF antigens represented a discontinuous epitope (see original paper I).

Lately peptide mimotopes have been suggested for cancer vaccine development as they can have the immunogenic potential to elicit stronger recognition of cancer antigens of both, protein and non-protein nature like the tumour-associated carbohydrate antigens GD2 disialoganglioside and CEACAM-5 glycoprotein, as well as provide a compact immunisation agent that can be coupled to additional stimulants of immune response[84]. However, to consider the use of a mimotope instead of its corresponding epitope for therapeutic applications it is also necessary to show that this peptide can induce the formation of the antibodies recognising the initial protein and not just bind the same antibody as the antibody-binding pockets may be different for both peptides and hence the immunisation may not yield the response recognising the mimicked antigen [149]. There were several unnatural peptide antigens in our collection that were recognised by cancer sera antibodies with statistical significance (Zayakin P et al., article in preparation) and there might be potentially useful therapeutic mimotopes among them.

The natural ORF antigen category comprised 130 non-redundant antigens among which 46 were well-characterized CT antigens like the members of CTAG, MAGE etc. families, 5 were autoantigens such as ANXA11 [150], and AKAP12 [151] known to induce antibody production in autoimmune disorders, 20 were antigens that have been previously identified by conventional SEREX, 34 antigens were from protein families whose other members have been detected by SEREX, and 25 were novel antigens against which no immune response has been reported before. The analysis of the natural ORF antigen clone sequences revealed that all of them were wild-type proteins and no mutations were found, but uncharacterised alternative splice variants (ASVs) were detected for three antigens. Also it is not excluded that mutations, splicing defects, chromosomal translocation or altered PTMs outside the cloned region could have caused structural changes revealing otherwise hidden regions to the antigen presentation machinery leading to the immune recognition of an otherwise tolerated protein[149], however this was not studied in the current work. Another reason of the immunogenicity of a wild-type protein is its ectopic expression or overrepresentation in cancer. To identify novel potential immunotherapy targets we selected all uncharacterised antigens and subjected these to a number of prioritization criteria in order to select the most interesting candidates for mRNA expression analyses. First, antigens recognised by cancer-related autoantibodies were identified by using our custom antigen microarray comprising 1158 antigens and non-recombinant phage controls for screening of sera from 190 melanoma, 173 gastric, and 13 breast cancer patients as well as 153 healthy donors. 50 natural ORF antigens showing cancer-related autoantibodies were further prioritised by the autoantibody frequency in cancer patients, their EST profile comparing normal to tumour and germ cell tissues, their putative involvement in oncogenesis and their structural motifs, putting a negative mark on coiled-coil domains, which are known as structures often recognised by cross-reactive antibodies, and a positive mark on surface localised antigens as they allow the development of additional antibody-based therapeutics. 23 top rated natural ORF antigens and five cancer-related undefined ORF antigens representing novel genes were subjected to mRNA expression analysis in various normal tissues and tumour-normal tissue pairs of

melanoma, gastric and breast cancer patients to determine whether the altered expression level might be the reason of immune recognition in cancer and to discern possible therapeutic targets (see original paper II).

We report here for the first time the connection of LRRC50 to cancer, and have shown it as a frequent breast and melanoma cancer antigen. Previous studies analysing patients with *situs inversus* and reduced muco-ciliary clearance of the respiratory tract have identified truncating mutation in LRRC50[152] what relates it to the proper functions of motile cilia during both, embryogenesis and adulthood [153]. A closely related structure – the non-motile primary cilium has been shown to be crucial for proper signalling during embryogenesis and adult tissue homeostasis through the Hedgehog, PDGFRαα and Wnt pathways [154], which are also deregulated in various cancers[155-157]. As there are no motile cilia on most cancer cells it would be interesting to determine if the overexpression of LRRC50 in tumour cells can participate in the signalling pathways of the primary cilium that promote the malignant phenotype. Our initial studies also indicate to the participation of LRRC50 in mitosis (unpublished observations), which might suggest another possible link to oncogenesis. The second novel CT antigen ESCO1 is an acetyltransferase involved in sister chromatid cohesion and DNA repair [158], and if its expression or function is deregulated it may contribute to the genetic instability [159]. We have demonstrated for the first time a cancer-associated immune response against ESCO1 with significant overexpression in melanomas. Further analyses of protein expression in cancer and normal tissues and the oncogenic potential of these genes will allow to assess their suitability for immunotherapy.

Cancer-related autoantibody response and testis-associated expression among normal tissues were detected for four annotated genes with unknown functions (ALLC, C11ORF20, C16ORF82, and C21ORF66) and four novel genes (LOC284861, LOC392843, clone #232, and clone #200). The expression of the testis-selective genes, ALLC, C11ORF20, C16ORF82, and LOC284861, were also upregulated in some tumour samples and these might potentially be novel members of the CT antigen group. The testis-restricted transcripts, C21ORF66, LOC392843, clone #232, and clone #200, were not detected in any analysed tumour samples, however they have been recognised by tumour autoantibodies and a larger sample size or different and more sensitive approaches directed to analysis of cancer stem cell population should be used to determine if they might represent genes expressed by seldom cancer stem cells as the germ cell-specific genes have been suggested to be involved in the control and establishment of stemness[28,90].

The cancer-related antigens ACTR2, SPARC, LIG1 and NOL8 have all been previously reported with potential implications in oncogenesis including cell motility[160], metastasis through affecting extra-cellular matrix[161,162], genetic instability[163], and cellular growth[164], respectively. We have shown that these genes are overexpressed in cancer, however the mRNA level in the overexpressed tumours didn't exceed the level of other normal tissue types. We suspect that the immunogenicity of these proteins is defined by other molecular alterations, and they cannot be suitable immunotherapy targets unless cancer-specific epitopes can be identified.

Among the 50 cancer-related natural ORF antigens there was also SPAG8, a sperm surface protein and a member of the sperm associated antigen group. SPAG proteins are expressed in germ cells and can induce immune response leading to infertility. Recently, implications in oncogenesis have also been demonstrated for several SPAG proteins, hence we decided to characterise this group of proteins to determine if there might be novel CT antigens among them with possible application in immunotherapy. The mRNA expression analysis showed that out of the 15 members only five genes SPAG1, SPAG6, SPAG8, SPAG15, SPAG17 and distinct isoforms of SPAG16 (SPAG16-L) were testis-associated.

All of these except SPAG15 were also upregulated in tumour tissues and cancer-related autoantibodies were found for SPAG1, SPAG6, SPAG8 and a testis-restricted splice variant of SPAG17 by screening a custom SPAG antigen microarray with sera from 543 patients with various cancers assigning these as novel CT antigens. SPAG6, SPAG16-L and SPAG17 have been shown to be necessary for the motility of sperm flagella[165,166] and other motile cilia of the human body, including epithelia of the brain and respiratory tract. [167,168]. Similarly to LRRC50 it would be of interest to determine if these proteins may participate in the non-motile cilium-mediated signalling and tumourigenic events. SPAG8 has been shown to participate in the regulation of mitosis[169], which could link it to the possible involvement in oncogenesis. In normal tissues this protein was detected only in rare cells of the parabasal layer of ectocervix and the glandular epithelia of the breast and stomach, and considering the frequency of the positive cells, they might be the stem cells of these epithelia providing additional ground for the therapeutic potential of this antigen (see original paper III).

We and also other authors have suggested that deregulated alternative splicing in cancer can lead to the producing untolerised protein isoforms[170-172] (see original paper IV). However, strictly cancer-specific splice variants generated by the recognition of the so-called cryptic splice sites have been rarely reported in connection to their immunogenicity. One such example is demonstrated for the gene, *HMSD*, that possesses an intronic SNP in a consensus splice site leading to a novel immunogenic splice variant[173]. The data of alternative transcripts deposited in the AceView database[174] as well as our previous bioinformatics analysis (unpublished data) suggest that such immunoprivileged tissues as testis and brain posses the largest variety of alternative transcripts among other normal tissues, and we proposed that it could be more often that alternative splicing deregulation created immunogenic protein isoforms in cancer due to ectopic recognition of such tissue-restricted splice sites. To evaluate the possibility of alternative splicing deregulation to be the cause of the immunogenicity of antigens representing novel ASVs we also analysed their mRNA expression in various normal and tumour-normal tissue pairs. We showed that novel splice variants of CFL1 and COX6B2 were testis-selective and also elevated in some cancer samples, while the novel transcript of TEF that was frequently recognised by cancer antibodies and the novel ASV of EVI5L were not detected in any sample, and further experiments are necessary to relate the deregulation of alternative splicing to the immunogenicity of these antigens. Another novel testis-restricted splicing isoform with overexpression in cancer was identified for SPAG17, and for the first time we demonstrated a testis-restricted splice variant-specific humoral immune response in cancer patients. We suggest that such antigens are designated to a separate category called CT-spliced antigens, as the mechanism of their production is different from the CT antigens namely alternative splicing errors rather than transcription deregulation. However their potential for immunotherapy application needs to be studied further as their expression can be very heterogeneous in the same tumour dictated by the splicing events in each cell that are not inherited across cell divisions and they are met in lower level simply because they represent only a fraction of the total mRNA pool of the corresponding gene, nevertheless they can provide novel tumor-specific targeting approaches with defined epitopes as well as might represent functionally important protein alterations in cancer.

To assess naturally occurring T cell responses against two immunotherapy candidate antigens LRRC50 and SPAG8 we chose to apply one of the most sensitive and commonly used T cell activation methods - the ELISPOT assay. There are many different protocols for the activation and culturing of T cells prior to ELISPOT as well as for the ELISPOT procedure itself that can lead to different results among various laboratories and poor reproducibility. We participated in an international collaboration project, NEUCAPS,

which was aimed to determine the main reasons of the variability of obtained ELISPOT results among several laboratories and to develop a standardised, reproducible and robust ELISPOT assay (see original paper V). The protocol developed as a result of this project was adapted for our T cell analysis. However, to detect rare tumour antigen-specific T cells, it is necessary to propagate them prior to ELISPOT assay by at least 10 days of culture in the presence of the target antigen. In order to determine the best conditions for the T cell pre-sensitisation, we optimised the number of peptide additions used for T cell activation and found that a single stimulation gave the highest number of reacting cells. This setting was suitable also for the detection of NY-ESO-1-specific CD8+ T cells. Six breast and six gastric cancer patients were analysed for the presence of naturally occurring CTLs against LRRC50 and SPAG8, and SPAG8-specific CTL response was detected in one gastric cancer patient. However, the population of naïve CTLs was not subtracted prior to CD8+ T cell stimulation, thereto the mRNA level of SPAG8 in the tumour tissue of the gastric cancer patient didn't exceed the level in the adjacent normal tissues raising suspicion that naïve CTL activation was induced. The stimulation should be repeated with a RA45+ depleted CD8+ T cell population to draw certain conclusions, nevertheless this result shows that the pool of the predicted SPAG8 peptides contains some that can indeed bind to the MHCI molecules and that SPAG8 is capable to induce a CTL response.

In order to successfully manipulate adoptive immune response one has to decipher the extremely complex, mutually regulating interactions among a large variety of B, T and DC cell subsets together with other not less diverse innate immune cells, and this has been one of the most challenging and difficult tasks in modern biology and medicine. It has to be kept in mind that most of the immune regulatory processes are still in the "model" phase, and these models are constantly being challenged from various sides and at various levels. For example, a recent publication revealed that the migration of naïve CD8+ T cells to tumour sites takes place, thereto these cells could be activated *in situ* that was independent of draining lymph nodes or even antigen presenting cells showing that tumours can support the differentiation of antigen-specific CTL effectors[175] circumventing the "mechanical" tolerance aspect, the "Danger Model", as well as the three signal activation paradigm. Besides, many "black holes" still exist in our understanding on (i) the role of various recently described T helper cell subsets in the anti-tumour immune response, on (ii) B cell activation and antibody production against cancer antigens, including the interactions with various T helper cell subsets, as well as on (iii) the regulation of CTL activation and T cell tolerance including antigen-specific T regulatory cell inhibition, and many more, which constitute our future scientific interests. Although only rare success stories of adaptive immunotherapy have been met, now at least we realise what we don't understand and can start to address these issues marking the entrance into the finishing straight of the marathon for the "right antigen", "right adjuvant", and " right immune response"[106,176].

5 Conclusions

- A large majority of 1328 serum-reactive recombinant phage clones represented antigens translated as unnatural ORF peptides - 83%, followed by wild-type antigens - 15%, and 2% constituted undefined ORF clones comprising putative novel genes and uncharacterised novel splice variants.
- Cancer patients' autoantibodies recognise the unnatural ORF peptides and not their natural ORF counterparts.
- Large-scale serum profiling of the antigen collection revealed cancer-related humoral immune response against 50 novel wild-type antigens and 8 undefined ORF antigens.
- The elaborated antigen prioritization scheme allowed the selection of 30 most promising candidate antigens for a comprehensive gene expression analysis to identify novel potential immunotherapy targets.
- mRNA expression analysis in various normal tissues revealed (1) six testis-selective genes (predominantly expressed in germ cells and placenta): ALLC, C16ORF82, C11ORF20, ESCO1, LRRC50, and the uncharacterised gene LOC284861, (2) two novel testis-selective alternative splice variants for CFL1 and COX6B2, and (3) four testis-restricted genes (exclusively expressed in germ cells or placenta): C21ORF66, uncharacterised genes LOC392843, clones #200 and #232.
- mRNA expression analysis in various cancer tissues revealed two novel cancer-testis (CT) antigens with potential application in cancer immunotherapy: LRRC50 was overexpressed in 23%-60% of various cancers, and ESCO1 – in 53% of melanoma samples.
- Among the 30 candidate antigens 16 were ubiquitous wild-type self-proteins and the reasons of immunogenicity of these antigens in cancer patients remain elusive.
- Expression analysis of the 15 members of the sperm-associated antigen group revealed five testis-selective genes: SPAG1, SPAG6, SPAG8, SPAG15, and SPAG17; four of them were identified as novel CT antigens: SPAG1, SPAG6, SPAG8, and SPAG17, which may have a potential application in tumour immunotherapy.
- A specific humoral immune response in cancer patients was demonstrated against a testis-restricted alternative splicing isoform of SPAG17 that, together with other immunogenic alternative splicing isoforms shared between germ cells and cancer, can be ascribed to a novel group of cancer antigens called CT-spliced antigens.
- The elaborated protocol of T cell activation analysis including CD8+ and CD4+ T cell separation from peripheral blood, *in vitro* pre-sensitisation and ELISPOT was capable to detect tumour-antigen specific CTL responses.
- SPAG8-specific naturally occurring CTLs were detected in one gastric cancer patient suggesting that therapy-induced immune recognition of this antigen could be possible.

Main thesis of defence

I The exploration of humoral immune response in cancer patients can reveal clinically significant tumour antigens.

II The group of SPAG proteins contains ubiquitous genes and genes predominantly expressed in testis that are significantly upregulated in cancer, elicit humoral immune response and can be qualified as novel CT antigens with potential implications in tumour immunotherapy.

III The pattern of alternative splicing variants partially overlaps in tumours and germ cells that contributes to the generation of immunologically untolerised protein isoforms in cancer cells.

Acknowledgements

This study was supported by:

EU 6th Framework Program ENACT contract No. LSHC-CT-2004-503306,
Grant from Latvian State Research Program No. 07-VP-2,
Grants from Latvian Council of Science No. 05.1623, 09.1288
European Social Fund project No. 1DP/1.1.2.1.2/09/IPIA/VIAA/004,
K. Morbergs Scholarship, University of Latvia
L'Oreal Scholarship for Women in Science.

I am sincerely thankful to:

My supervisor
Dr. Aija Linē for everything that I know and have learned.

My colleagues from Latvian Biomedical Research and Study Centre
Dr. Zane Kalniņa, Pavel Zayakin, Artūrs Ābols, Lāsma Ivanova, Elīna Zandberga, Dr. Ēriks Jankevics, Dr. Inna Iņaškina, Dr. Jekaterina Ērenpreisa, Ludmila Jackeviča, Ramona Petrovska, Dr. Ruta Brūvere, Dr. Dace Pjanova,
as well as colleagues from the Latvian Institute of Organic Synthesis
Dr. Edgars Liepiņš, Dr. Maija Dambrova, Dr. Irina Šestakova for their priceless help in everyday work.

Local and international collaborators:
Dr. Aivars Stengrēvics from Riga Eastern Clinical University Hospital,
Dr. Liene Ņikitina-Zaķe from the Genome Database of the Latvian population,
Dr. Kristīne Ducena and Dr. Viktors Kozirovskis from Pauls Stradins Clinical University Hospital, Riga,
Dr. Dirk Schadendorf from Skin Cancer Unit in German Cancer Research Center,
Dr. Thomas Wex from Clinic of Gastroenterology, Hepatology and Infectious Diseases, Otto-von-Guericke University Magdeburg,
Dr. Gustav Gaudernak from Norwegian Radium Hospital and
Onyvax Vaccine Therapies Ltd, UK for the clinical material.

Prof. Alexander Knuth, Dr. Maries van den Broek, Natko Nuber, and Stefanie Dannenman from the University Hospital Zurich, Division of Oncology, Dr. Sacha Gnjatic from Ludwig Institute for Cancer Research Ltd, Memorial-Sloan Kettering Cancer Center, New York, and Dr. Stéphanie McArdle from School of Science and Technology, Nottingham Trent University for their excellent advice, consultations and help in the development of T cell activation assay.

Dr. Arturs Šorubalko and Jurgita Priedīte from Pauls Stradins Clinical University Hospital, Riga, Radiotherapy Division for their kind help in cell irradiation.

And my family and friends for the truly unlimited mental as well as nutritional (mom, I love those meat balls!) support.

References

1. Buckwalter MR and Srivastava PK. "It is the antigen(s), stupid" and other lessons from over a decade of vaccitherapy of human cancer. Semin Immunol. 2008;20(5):296-300.
2. Finn OJ. Cancer immunology. N Engl J Med. 2008;358(25):2704-2715.
3. Rosenberg SA. Development of effective immunotherapy for the treatment of patients with cancer. J Am Coll Surg. 2004;198(5):685-696.
4. Baxevanis CN, Perez SA, and Papamichail M. Combinatorial treatments including vaccines, chemotherapy and monoclonal antibodies for cancer therapy. Cancer Immunol Immunother. 2009;58(3):317-324.
5. Zhang T and Herlyn D. Combination of active specific immunotherapy or adoptive antibody or lymphocyte immunotherapy with chemotherapy in the treatment of cancer. Cancer Immunol Immunother. 2009;58(4):475-492.
6. Cheever MA, Allison JP, Ferris AS, et al. The prioritization of cancer antigens: a national cancer institute pilot project for the acceleration of translational research. Clin Cancer Res. 2009;15(17):5323-5337.
7. Flajnik MF and Kasahara M. Origin and evolution of the adaptive immune system: genetic events and selective pressures. Nat Rev Genet. 2010;11(1):47-59.
8. Kamradt T and Mitchison NA. Tolerance and autoimmunity. N Engl J Med. 2001;344(9):655-664.
9. Hogquist KA, Baldwin TA, and Jameson SC. Central tolerance: learning self-control in the thymus. Nat Rev Immunol. 2005;5(10):772-782.
10. von Boehmer H and Melchers F. Checkpoints in lymphocyte development and autoimmune disease. Nat Immunol. 2010;11(1):14-20.
11. Melchers F and Rolink AR. B cell tolerance--how to make it and how to break it. Curr Top Microbiol Immunol. 2006;305:1-23.
12. Meffre E and Wardemann H. B-cell tolerance checkpoints in health and autoimmunity. Curr Opin Immunol. 2008;20(6):632-638.
13. Mueller DL. Mechanisms maintaining peripheral tolerance. Nat Immunol. 2010;11(1):21-27.
14. Stadanlick JE and Cancro MP. BAFF and the plasticity of peripheral B cell tolerance. Curr Opin Immunol. 2008;20(2):158-161.
15. Lim HW, Hillsamer P, Banham AH, et al. Cutting edge: direct suppression of B cells by CD4+ CD25+ regulatory T cells. J Immunol. 2005;175(7):4180-4183.
16. Lim HW, Hillsamer P, and Kim CH. Regulatory T cells can migrate to follicles upon T cell activation and suppress GC-Th cells and GC-Th cell-driven B cell responses. J Clin Invest. 2004;114(11):1640-1649.
17. Matzinger P. The danger model: a renewed sense of self. Science. 2002;296(5566):301-305.
18. Eisenlohr LC and Rothstein JL. Antigen processing and presentation. Cancer Treat Res. 2005;123:3-36.
19. Quah BJ and O'Neill HC. Maturation of function in dendritic cells for tolerance and immunity. J Cell Mol Med. 2005;9(3):643-654.
20. Cools N, Ponsaerts P, Van Tendeloo VF, et al. Balancing between immunity and tolerance: an interplay between dendritic cells, regulatory T cells, and effector T cells. J Leukoc Biol. 2007;82(6):1365-1374.
21. Vignali D. How many mechanisms do regulatory T cells need? Eur J Immunol. 2008;38(4):908-911.

22. Fletcher AL, Lukacs-Kornek V, Reynoso ED, et al. Lymph node fibroblastic reticular cells directly present peripheral tissue antigen under steady-state and inflammatory conditions. J Exp Med. 2010;207(4):689-697.
23. Cohen JN, Guidi CJ, Tewalt EF, et al. Lymph node-resident lymphatic endothelial cells mediate peripheral tolerance via Aire-independent direct antigen presentation. J Exp Med. 2010;207(4):681-688.
24. Reynoso ED, Lee JW, and Turley SJ. Peripheral tolerance induction by lymph node stroma. Adv Exp Med Biol. 2009;633:113-127.
25. Zhang J, Xu X, and Liu Y. Activation-induced cell death in T cells and autoimmunity. Cell Mol Immunol. 2004;1(3):186-192.
26. Srinivasan M and Frauwirth KA. Peripheral tolerance in CD8+ T cells. Cytokine. 2009;46(2):147-159.
27. Mescher MF, Agarwal P, Casey KA, et al. Molecular basis for checkpoints in the CD8 T cell response: tolerance versus activation. Semin Immunol. 2007;19(3):153-161.
28. Yang XF. Immunology of stem cells and cancer stem cells. Cell Mol Immunol. 2007;4(3):161-171.
29. Bouaziz JD, Yanaba K, and Tedder TF. Regulatory B cells as inhibitors of immune responses and inflammation. Immunol Rev. 2008;224:201-214.
30. Suzuki M, Konya C, Goronzy JJ, et al. Inhibitory CD8+ T cells in autoimmune disease. Hum Immunol. 2008;69(11):781-789.
31. Beissert S, Schwarz A, and Schwarz T. Regulatory T cells. J Invest Dermatol. 2006;126(1):15-24.
32. Smith TR and Kumar V. Revival of CD8+ Treg-mediated suppression. Trends Immunol. 2008;29(7):337-342.
33. Jiang H and Chess L. How the immune system achieves self-nonself discrimination during adaptive immunity. Adv Immunol. 2009;102:95-133.
34. Sadegh-Nasseri S, Dalai SK, Korb Ferris LC, et al. Suboptimal engagement of the T-cell receptor by a variety of peptide-MHC ligands triggers T-cell anergy. Immunology. 2010;129(1):1-7.
35. Mueller SN and Germain RN. Stromal cell contributions to the homeostasis and functionality of the immune system. Nat Rev Immunol. 2009;9(9):618-629.
36. Zhou L, Chong MM, and Littman DR. Plasticity of CD4+ T cell lineage differentiation. Immunity. 2009;30(5):646-655.
37. Wilson CB, Rowell E, and Sekimata M. Epigenetic control of T-helper-cell differentiation. Nat Rev Immunol. 2009;9(2):91-105.
38. Martorelli D, Muraro E, Merlo A, et al. Role of CD4+ cytotoxic T lymphocytes in the control of viral diseases and cancer. Int Rev Immunol. 2010;29(4):371-402.
39. Zhu J and Paul WE. CD4 T cells: fates, functions, and faults. Blood. 2008;112(5):1557-1569.
40. Annunziato F and Romagnani S. Heterogeneity of human effector CD4+ T cells. Arthritis Res Ther. 2009;11(6):257.
41. King C, Tangye SG, and Mackay CR. T follicular helper (TFH) cells in normal and dysregulated immune responses. Annu Rev Immunol. 2008;26:741-766.
42. Gowthaman U, Chodisetti SB, and Agrewala JN. T cell help to B cells in germinal centers: putting the jigsaw together. Int Rev Immunol.29(4):403-420.
43. Yu D and Vinuesa CG. The elusive identity of T follicular helper cells. Trends Immunol. 2010(10):377-83.
44. Park H, Li Z, Yang XO, et al. A distinct lineage of CD4 T cells regulates tissue inflammation by producing interleukin 17. Nat Immunol. 2005;6(11):1133-1141.

45. Zou W and Restifo NP. T(H)17 cells in tumour immunity and immunotherapy. Nat Rev Immunol. 2010;10(4):248-256.
46. Soroosh P and Doherty TA. Th9 and allergic disease. Immunology. 2009;127(4):450-458.
47. van den Ham HJ, de Waal L, Andeweg AC, et al. Identification of helper T cell master regulator candidates using the polar score method. J Immunol Methods. 2010.
48. Staudt V, Bothur E, Klein M, et al. Interferon-regulatory factor 4 is essential for the developmental program of T helper 9 cells. Immunity. 2010;33(2):192-202.
49. Trifari S, Kaplan CD, Tran EH, et al. Identification of a human helper T cell population that has abundant production of interleukin 22 and is distinct from T(H)-17, T(H)1 and T(H)2 cells. Nat Immunol. 2009;10(8):864-871.
50. Duhen T, Geiger R, Jarrossay D, et al. Production of interleukin 22 but not interleukin 17 by a subset of human skin-homing memory T cells. Nat Immunol. 2009;10(8):857-863.
51. Fujita H, Nograles KE, Kikuchi T, et al. Human Langerhans cells induce distinct IL-22-producing CD4+ T cells lacking IL-17 production. Proc Natl Acad Sci U S A. 2009;106(51):21795-21800.
52. Eyerich S, Eyerich K, Pennino D, et al. Th22 cells represent a distinct human T cell subset involved in epidermal immunity and remodeling. J Clin Invest. 2009;119(12):3573-3585.
53. Trifari S and Spits H. IL-22-producing CD4+ T cells: middle-men between the immune system and its environment. Eur J Immunol. 2010;40(9):2369-2371.
54. Ramirez JM, Brembilla NC, Sorg O, et al. Activation of the aryl hydrocarbon receptor reveals distinct requirements for IL-22 and IL-17 production by human T helper cells. Eur J Immunol.40(9):2450-2459.
55. Vinuesa CG, Tangye SG, Moser B, et al. Follicular B helper T cells in antibody responses and autoimmunity. Nat Rev Immunol. 2005;5(11):853-865.
56. Casey KA and Mescher MF. IL-21 promotes differentiation of naive CD8 T cells to a unique effector phenotype. J Immunol. 2007;178(12):7640-7648.
57. Andersen MH, Schrama D, Thor Straten P, et al. Cytotoxic T cells. J Invest Dermatol. 2006;126(1):32-41.
58. Croft M, Carter L, Swain SL, et al. Generation of polarized antigen-specific CD8 effector populations: reciprocal action of interleukin (IL)-4 and IL-12 in promoting type 2 versus type 1 cytokine profiles. J Exp Med. 1994;180(5):1715-1728.
59. Vizler C, Bercovici N, Heurtier A, et al. Relative diabetogenic properties of islet-specific Tc1 and Tc2 cells in immunocompetent hosts. J Immunol. 2000;165(11):6314-6321.
60. Dobrzanski MJ, Reome JB, and Dutton RW. Therapeutic effects of tumor-reactive type 1 and type 2 CD8+ T cell subpopulations in established pulmonary metastases. J Immunol. 1999;162(11):6671-6680.
61. Gershon RK and Kondo K. Cell interactions in the induction of tolerance: the role of thymic lymphocytes. Immunology. 1970;18(5):723-737.
62. Hamada H, Garcia-Hernandez Mde L, Reome JB, et al. Tc17, a unique subset of CD8 T cells that can protect against lethal influenza challenge. J Immunol. 2009;182(6):3469-3481.
63. Kondo T, Takata H, Matsuki F, et al. Cutting edge: Phenotypic characterization and differentiation of human CD8+ T cells producing IL-17. J Immunol. 2009;182(4):1794-1798.

64. McHeyzer-Williams LJ, Malherbe LP, and McHeyzer-Williams MG. Checkpoints in memory B-cell evolution. Immunol Rev. 2006;211:255-268.
65. Shapiro-Shelef M and Calame K. Regulation of plasma-cell development. Nat Rev Immunol. 2005;5(3):230-242.
66. MacLennan IC, Toellner KM, Cunningham AF, et al. Extrafollicular antibody responses. Immunol Rev. 2003;194:8-18.
67. Batista FD and Harwood NE. The who, how and where of antigen presentation to B cells. Nat Rev Immunol. 2009;9(1):15-27.
68. Harris DP, Haynes L, Sayles PC, et al. Reciprocal regulation of polarized cytokine production by effector B and T cells. Nat Immunol. 2000;1(6):475-482.
69. Chen X and Jensen PE. The role of B lymphocytes as antigen-presenting cells. Arch Immunol Ther Exp (Warsz). 2008;56(2):77-83.
70. Nimmerjahn F and Ravetch JV. Antibodies, Fc receptors and cancer. Curr Opin Immunol. 2007;19(2):239-245.
71. Carter P. Improving the efficacy of antibody-based cancer therapies. Nat Rev Cancer. 2001;1(2):118-129.
72. Ullrich E, Bonmort M, Mignot G, et al. Tumor stress, cell death and the ensuing immune response. Cell Death Differ. 2008;15(1):21-28.
73. de Visser KE, Eichten A, and Coussens LM. Paradoxical roles of the immune system during cancer development. Nat Rev Cancer. 2006;6(1):24-37.
74. Dunn GP, Old LJ, and Schreiber RD. The immunobiology of cancer immunosurveillance and immunoediting. Immunity. 2004;21(2):137-148.
75. Reiman JM, Kmieciak M, Manjili MH, et al. Tumor immunoediting and immunosculpting pathways to cancer progression. Semin Cancer Biol. 2007;17(4):275-287.
76. Ellyard JI, Simson L, and Parish CR. Th2-mediated anti-tumour immunity: friend or foe? Tissue Antigens. 2007;70(1):1-11.
77. Kim R, Emi M, and Tanabe K. Cancer immunoediting from immune surveillance to immune escape. Immunology. 2007;121(1):1-14.
78. Prehn RT and Main JM. Immunity to methylcholanthrene-induced sarcomas. J Natl Cancer Inst. 1957;18(6):769-778.
79. van der Bruggen P, Traversari C, Chomez P, et al. A gene encoding an antigen recognized by cytolytic T lymphocytes on a human melanoma. Science. 1991;254(5038):1643-1647.
80. Yang F and Yang XF. New concepts in tumor antigens: their significance in future immunotherapies for tumors. Cell Mol Immunol. 2005;2(5):331-341.
81. Vollmers HP and Brandlein S. Tumors: too sweet to remember? Mol Cancer. 2007;6:78.
82. Vyas JM, Van der Veen AG, and Ploegh HL. The known unknowns of antigen processing and presentation. Nat Rev Immunol. 2008;8(8):607-618.
83. Fuster MM and Esko JD. The sweet and sour of cancer: glycans as novel therapeutic targets. Nat Rev Cancer. 2005;5(7):526-542.
84. Knittelfelder R, Riemer AB, and Jensen-Jarolim E. Mimotope vaccination--from allergy to cancer. Expert Opin Biol Ther. 2009;9(4):493-506.
85. Finn OJ. Human tumor antigens, immunosurveillance, and cancer vaccines. Immunol Res. 2006;36(1-3):73-82.
86. Schietinger A, Philip M, and Schreiber H. Specificity in cancer immunotherapy. Semin Immunol. 2008;20(5):276-285.

87. Hanada K, Yewdell JW, and Yang JC. Immune recognition of a human renal cancer antigen through post-translational protein splicing. Nature. 2004;427(6971):252-256.
88. Ng B, Yang F, Huston DP, et al. Increased noncanonical splicing of autoantigen transcripts provides the structural basis for expression of untolerized epitopes. J Allergy Clin Immunol. 2004;114(6):1463-1470.
89. Gnjatic S, Nishikawa H, Jungbluth AA, et al. NY-ESO-1: review of an immunogenic tumor antigen. Adv Cancer Res. 2006;95:1-30.
90. Simpson AJ, Caballero OL, Jungbluth A, et al. Cancer/testis antigens, gametogenesis and cancer. Nat Rev Cancer. 2005;5(8):615-625.
91. Hanahan D and Weinberg RA. The hallmarks of cancer. Cell. 2000;100(1):57-70.
92. Zitvogel L, Tesniere A, and Kroemer G. Cancer despite immunosurveillance: immunoselection and immunosubversion. Nat Rev Immunol. 2006;6(10):715-727.
93. Bai XF, Liu J, Li O, et al. Antigenic drift as a mechanism for tumor evasion of destruction by cytolytic T lymphocytes. J Clin Invest. 2003;111(10):1487-1496.
94. Khong HT and Restifo NP. Natural selection of tumor variants in the generation of "tumor escape" phenotypes. Nat Immunol. 2002;3(11):999-1005.
95. Beyer M and Schultze JL. Regulatory T cells: major players in the tumor microenvironment. Curr Pharm Des. 2009;15(16):1879-1892.
96. Ostrand-Rosenberg S and Sinha P. Myeloid-derived suppressor cells: linking inflammation and cancer. J Immunol. 2009;182(8):4499-4506.
97. Valenti R, Huber V, Filipazzi P, et al. Human tumor-released microvesicles promote the differentiation of myeloid cells with transforming growth factor-beta-mediated suppressive activity on T lymphocytes. Cancer Res. 2006;66(18):9290-9298.
98. Dunn GP, Koebel CM, and Schreiber RD. Interferons, immunity and cancer immunoediting. Nat Rev Immunol. 2006;6(11):836-848.
99. Grivennikov SI, Greten FR, and Karin M. Immunity, inflammation, and cancer. Cell. 2010;140(6):883-899.
100. Ahmad M, Rees RC, and Ali SA. Escape from immunotherapy: possible mechanisms that influence tumor regression/progression. Cancer Immunol Immunother. 2004;53(10):844-854.
101. Seruga B, Zhang H, Bernstein LJ, et al. Cytokines and their relationship to the symptoms and outcome of cancer. Nat Rev Cancer. 2008;8(11):887-899.
102. Sengupta N, MacFie TS, MacDonald TT, et al. Cancer immunoediting and "spontaneous" tumor regression. Pathol Res Pract. 2010;206(1):1-8.
103. Johansson M, Denardo DG, and Coussens LM. Polarized immune responses differentially regulate cancer development. Immunol Rev. 2008;222:145-154.
104. Finn OJ. Cancer vaccines: between the idea and the reality. Nat Rev Immunol. 2003;3(8):630-641.
105. Orimo A and Weinberg RA. Stromal fibroblasts in cancer: a novel tumor-promoting cell type. Cell Cycle. 2006;5(15):1597-1601.
106. Pejawar-Gaddy S and Finn OJ. Cancer vaccines: accomplishments and challenges. Crit Rev Oncol Hematol. 2008;67(2):93-102.
107. Copier J and Dalgleish A. Whole-cell vaccines: A failure or a success waiting to happen? Curr Opin Mol Ther. 2010;12(1):14-20.
108. Segal BH, Wang XY, Dennis CG, et al. Heat shock proteins as vaccine adjuvants in infections and cancer. Drug Discov Today. 2006;11(11-12):534-540.
109. Osada T, Clay TM, Woo CY, et al. Dendritic cell-based immunotherapy. Int Rev Immunol. 2006;25(5-6):377-413.

110. Jager E, Jager D, and Knuth A. Clinical cancer vaccine trials. Curr Opin Immunol. 2002;14(2):178-182.
111. Rosenberg SA, Restifo NP, Yang JC, et al. Adoptive cell transfer: a clinical path to effective cancer immunotherapy. Nat Rev Cancer. 2008;8(4):299-308.
112. Rosenberg SA, Sherry RM, Morton KE, et al. Tumor progression can occur despite the induction of very high levels of self/tumor antigen-specific CD8+ T cells in patients with melanoma. J Immunol. 2005;175(9):6169-6176.
113. Zou W. Immunosuppressive networks in the tumour environment and their therapeutic relevance. Nat Rev Cancer. 2005;5(4):263-274.
114. Zou W. Regulatory T cells, tumour immunity and immunotherapy. Nat Rev Immunol. 2006;6(4):295-307.
115. Muller AJ and Scherle PA. Targeting the mechanisms of tumoral immune tolerance with small-molecule inhibitors. Nat Rev Cancer. 2006;6(8):613-625.
116. Schnurr M, Orban M, Robson NC, et al. ISCOMATRIX adjuvant induces efficient cross-presentation of tumor antigen by dendritic cells via rapid cytosolic antigen delivery and processing via tripeptidyl peptidase II. J Immunol. 2009;182(3):1253-1259.
117. Dubensky TW, Jr. and Reed SG. Adjuvants for cancer vaccines. Semin Immunol. 2010;22(3):155-161.
118. Zitvogel L, Apetoh L, Ghiringhelli F, et al. Immunological aspects of cancer chemotherapy. Nat Rev Immunol. 2008;8(1):59-73.
119. Gao JQ, Okada N, Mayumi T, et al. Immune cell recruitment and cell-based system for cancer therapy. Pharm Res. 2008;25(4):752-768.
120. Peggs KS, Quezada SA, and Allison JP. Cancer immunotherapy: co-stimulatory agonists and co-inhibitory antagonists. Clin Exp Immunol. 2009;157(1):9-19.
121. Gould DJ and Chernajovsky Y. Novel delivery methods to achieve immunomodulation. Curr Opin Pharmacol. 2007;7(4):445-450.
122. Dudley ME, Yang JC, Sherry R, et al. Adoptive cell therapy for patients with metastatic melanoma: evaluation of intensive myeloablative chemoradiation preparative regimens. J Clin Oncol. 2008;26(32):5233-5239.
123. Bot A. The landmark approval of Provenge, what it means to immunology and "in this issue": the complex relation between vaccines and autoimmunity. Int Rev Immunol. 2010;29(3):235-238.
124. Berry LJ, Moeller M, and Darcy PK. Adoptive immunotherapy for cancer: the next generation of gene-engineered immune cells. Tissue Antigens. 2009;74(4):277-289.
125. Eshhar Z. Adoptive cancer immunotherapy using genetically engineered designer T-cells: First steps into the clinic. Curr Opin Mol Ther. 2010;12(1):55-63.
126. Wu TY, Hsieh CC, Hong JJ, et al. IRSS: a web-based tool for automatic layout and analysis of IRES secondary structure prediction and searching system in silico. BMC Bioinformatics. 2009;10:160.
127. Roitt I, Brostoff J, and Male D, *Immunology*. Sixth Edition ed. 2001: Elsevier Science Ltd. 480.
128. Zinkernagel RM and Hengartner H. Regulation of the immune response by antigen. Science. 2001;293(5528):251-253.
129. Nishimura F, Dusak JE, Eguchi J, et al. Adoptive transfer of type 1 CTL mediates effective anti-central nervous system tumor response: critical roles of IFN-inducible protein-10. Cancer Res. 2006;66(8):4478-4487.
130. Ye Z, Tang C, Xu S, et al. Type 1 CD8+ T cells are superior to type 2 CD8+ T cells in tumor immunotherapy due to their efficient cytotoxicity, prolonged survival and type 1 immune modulation. Cell Mol Immunol. 2007;4(4):277-285.

131. Muranski P and Restifo NP. Adoptive immunotherapy of cancer using CD4(+) T cells. Curr Opin Immunol. 2009;21(2):200-208.
132. van de Berg PJ, van Leeuwen EM, ten Berge IJ, et al. Cytotoxic human CD4(+) T cells. Curr Opin Immunol. 2008;20(3):339-343.
133. Locksley RM. Nine lives: plasticity among T helper cell subsets. J Exp Med. 2009;206(8):1643-1646.
134. Sondergaard H, Galsgaard ED, Bartholomaeussen M, et al. Intratumoral interleukin-21 increases antitumor immunity, tumor-infiltrating CD8+ T-cell density and activity, and enlarges draining lymph nodes. J Immunother. 2010;33(3):236-249.
135. Mills KH. Induction, function and regulation of IL-17-producing T cells. Eur J Immunol. 2008;38(10):2636-2649.
136. Dardalhon V, Awasthi A, Kwon H, et al. IL-4 inhibits TGF-beta-induced Foxp3+ T cells and, together with TGF-beta, generates IL-9+ IL-10+ Foxp3(-) effector T cells. Nat Immunol. 2008;9(12):1347-1355.
137. Lund FE and Randall TD. Effector and regulatory B cells: modulators of CD4(+) T cell immunity. Nat Rev Immunol. 2010;10(4):236-247.
138. Bayry J, Lacroix-Desmazes S, Kazatchkine MD, et al. Modulation of dendritic cell maturation and function by B lymphocytes. J Immunol. 2005;175(1):15-20.
139. Reuschenbach M, von Knebel Doeberitz M, and Wentzensen N. A systematic review of humoral immune responses against tumor antigens. Cancer Immunol Immunother. 2009;58(10):1535-1544.
140. Srivastava PK. Do human cancers express shared protective antigens? or the necessity of remembrance of things past. Semin Immunol. 1996;8(5):295-302.
141. Pardoll DM. Inducing autoimmune disease to treat cancer. Proc Natl Acad Sci U S A. 1999;96(10):5340-5342.
142. Overwijk WW and Restifo NP. Autoimmunity and the immunotherapy of cancer: targeting the "self" to destroy the "other". Crit Rev Immunol. 2000;20(6):433-450.
143. Chen YT, Scanlan MJ, Sahin U, 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):1914-1918.
144. Kalnina Z, Silina K, Meistere I, et al. Evaluation of T7 and lambda phage display systems for survey of autoantibody profiles in cancer patients. J Immunol Methods. 2008;334(1-2):37-50.
145. Chatterjee M and Tainsky MA. Non-traditional immunogens and their application to immunotherapy. Curr Opin Mol Ther. 2008;10(1):62-67.
146. Taft RJ, Pang KC, Mercer TR, et al. Non-coding RNAs: regulators of disease. J Pathol. 2010;220(2):126-139.
147. Yazgan O and Krebs JE. Noncoding but nonexpendable: transcriptional regulation by large noncoding RNA in eukaryotes. Biochem Cell Biol. 2007;85(4):484-496.
148. Meloen RH, Puijk WC, and Slootstra JW. Mimotopes: realization of an unlikely concept. J Mol Recognit. 2000;13(6):352-359.
149. Van Regenmortel MH. What is a B-cell epitope? Methods Mol Biol. 2009;524:3-20.
150. Misaki Y, Pruijn GJ, van der Kemp AW, et al. The 56K autoantigen is identical to human annexin XI. J Biol Chem. 1994;269(6):4240-4246.
151. Gordon T, Grove B, Loftus JC, et al. Molecular cloning and preliminary characterization of a novel cytoplasmic antigen recognized by myasthenia gravis sera. J Clin Invest. 1992;90(3):992-999.

152. Loges NT, Olbrich H, Becker-Heck A, et al. Deletions and point mutations of LRRC50 cause primary ciliary dyskinesia due to dynein arm defects. Am J Hum Genet. 2009;85(6):883-889.
153. Michaud EJ and Yoder BK. The primary cilium in cell signaling and cancer. Cancer Res. 2006;66(13):6463-6467.
154. Veland IR, Awan A, Pedersen LB, et al. Primary cilia and signaling pathways in mammalian development, health and disease. Nephron Physiol. 2009;111(3):p39-53.
155. Yu J, Ustach C, and Kim HR. Platelet-derived growth factor signaling and human cancer. J Biochem Mol Biol. 2003;36(1):49-59.
156. MacDonald BT, Tamai K, and He X. Wnt/beta-catenin signaling: components, mechanisms, and diseases. Dev Cell. 2009;17(1):9-26.
157. Wong SY, Seol AD, So PL, et al. Primary cilia can both mediate and suppress Hedgehog pathway-dependent tumorigenesis. Nat Med. 2009;15(9):1055-1061.
158. Hou F and Zou H. Two human orthologues of Eco1/Ctf7 acetyltransferases are both required for proper sister-chromatid cohesion. Mol Biol Cell. 2005;16(8):3908-3918.
159. Zhang J, Shi X, Li Y, et al. Acetylation of Smc3 by Eco1 is required for S phase sister chromatid cohesion in both human and yeast. Mol Cell. 2008;31(1):143-151.
160. Millard TH, Sharp SJ, and Machesky LM. Signalling to actin assembly via the WASP (Wiskott-Aldrich syndrome protein)-family proteins and the Arp2/3 complex. Biochem J. 2004;380(Pt 1):1-17.
161. Bradshaw AD, Graves DC, Motamed K, et al. SPARC-null mice exhibit increased adiposity without significant differences in overall body weight. Proc Natl Acad Sci U S A. 2003;100(10):6045-6050.
162. Bradshaw AD, Puolakkainen P, Dasgupta J, et al. SPARC-null mice display abnormalities in the dermis characterized by decreased collagen fibril diameter and reduced tensile strength. J Invest Dermatol. 2003;120(6):949-955.
163. Tomkinson AE and Mackey ZB. Structure and function of mammalian DNA ligases. Mutat Res. 1998;407(1):1-9.
164. Sekiguchi T, Todaka Y, Wang Y, et al. A novel human nucleolar protein, Nop132, binds to the G proteins, RRAG A/C/D. J Biol Chem. 2004;279(9):8343-8350.
165. Zhang Z, Sapiro R, Kapfhamer D, et al. A sperm-associated WD repeat protein orthologous to Chlamydomonas PF20 associates with Spag6, the mammalian orthologue of Chlamydomonas PF16. Mol Cell Biol. 2002;22(22):7993-8004.
166. Zhang Z, Jones BH, Tang W, et al. Dissecting the axoneme interactome: the mammalian orthologue of Chlamydomonas PF6 interacts with sperm-associated antigen 6, the mammalian orthologue of Chlamydomonas PF16. Mol Cell Proteomics. 2005;4(7):914-923.
167. Lonergan KM, Chari R, Deleeuw RJ, et al. Identification of novel lung genes in bronchial epithelium by serial analysis of gene expression. Am J Respir Cell Mol Biol. 2006;35(6):651-661.
168. Zhang Z, Tang W, Zhou R, et al. Accelerated mortality from hydrocephalus and pneumonia in mice with a combined deficiency of SPAG6 and SPAG16L reveals a functional interrelationship between the two central apparatus proteins. Cell Motil Cytoskeleton. 2007;64(5):360-376.
169. Li R, Tang XL, Miao SY, et al. Regulation of the G2/M phase of the cell cycle by sperm associated antigen 8 (SPAG8) protein. Cell Biochem Funct. 2009;27(5):264-268.

170. Yang F, Chen IH, Xiong Z, et al. Model of stimulation-responsive splicing and strategies in identification of immunogenic isoforms of tumor antigens and autoantigens. Clin Immunol. 2006;121(2):121-133.
171. Kalnina Z, Zayakin P, Silina K, et al. Alterations of pre-mRNA splicing in cancer. Genes Chromosomes Cancer. 2005;42(4):342-357.
172. Kalnina Z, Silina K, and Line A. Autoantibody Profiles as Biomarkers for Response to Therapy and Early Detection of Cancer. Curr Cancer Ther Rev. 2008;4(2):149-156.
173. Kawase T, Akatsuka Y, Torikai H, et al. Alternative splicing due to an intronic SNP in HMSD generates a novel minor histocompatibility antigen. Blood. 2007;110(3):1055-1063.
174. AceView, [database online]. Bethesda, MD: National Center for Biotechnology Information; 2000. Updated December 12, 2009.
175. Thompson ED, Enriquez HL, Fu YX, et al. Tumor masses support naive T cell infiltration, activation, and differentiation into effectors. J Exp Med. 2010;207(8):1791-1804.
176. Weiner LM. Cancer immunotherapy--the endgame begins. N Engl J Med. 2008;358(25):2664-2665.