Vaccine 31 (2013) 3805–3810

Contents lists available at SciVerse ScienceDirect

Vaccine

journal homepage: www.elsevier.com/locate/vaccine

Review

Proteins as T cell antigens: Methods for high-throughput identification

Daniel Grubaugh[a], Jessica Baker Flechtner[b,*], Darren E. Higgins[a,**]

[a] Department of Microbiology and Immunobiology, Harvard Medical School, 77 Avenue Louis Pasteur, Boston, MA 02115-5701, United States
[b] Genocea Biosciences, Inc., Cambridge Discovery Park, 100 Acorn Park Drive, 5th Floor, Cambridge MA 02142, United States

ARTICLE INFO

Article history:
Received 3 April 2013
Received in revised form 3 June 2013
Accepted 13 June 2013
Available online 24 June 2013

Keywords:
T cells
Protein antigen
Subunit vaccine
High-throughput screening
Reverse vaccinology, MHC

ABSTRACT

Vaccines are the most cost-effective means of preventing infectious diseases and have the potential to be used in a therapeutic capacity for the treatment of numerous chronic diseases and cancer. The majority of available vaccines function by eliciting antibodies that can neutralize toxins or opsonize the pathogen leading to elimination by professional phagocytes. However, there are many infectious and non-infectious diseases for which there are no available vaccines or the current antibody-mediated vaccines offer insufficient protection. There is emerging evidence that successful protection for these conditions requires the stimulation of T cell responses in addition to antibody. Genome/proteome-wide screening of pathogens to identify appropriate antibody targets for inclusion in vaccines has become widely used in recent years. However, the application of high-throughput proteomic screening approaches to identify T cell antigens has substantially lagged behind, primarily due to the lack of methods to identify full protein targets of T cell immunity across a broad human population. In this review, we will discuss some of the significant advances that have been made in high-throughput identification of T cell antigens for the development of novel efficacious vaccines.

Contents

1. Introduction

Advances in genomics have led to new approaches for vaccine development. The first organism with a fully sequenced genome was the pathogen *Haemophilus influenzae* in 1995 [1]. Since then, the genomes of thousands of strains of disease-causing pathogens have been sequenced. A notable translation of genomic information to yield novel vaccines is Reverse Vaccinology, which has led to new vaccines against group B *Streptococcus* and *Neisseria meningitidis* [2]. In Reverse Vaccinology, genomic data is used to predict proteins that are secreted or expressed on the pathogen surface, which are plausible targets for protective antibody responses. Based on these predictions, identified proteins are evaluated in animal models of disease to test *in vitro* correlates of immunity or *in vivo* protection. More recently, it has become possible to screen patient or animal immune sera against large segments of a pathogen's proteome to identify potential protective antigens [3]. Both of these techniques focus primarily on humoral (antibody-based) immunity. However, many pathogens and cancers are not controlled by humoral immunity alone. For these diseases, cellular immunity provided by $CD4^+$ and/or $CD8^+$ T cells will likely be a necessary component of any effective vaccination strategy [4–6]. The lack of known protective T cell antigens for many diseases or an efficient means to identify these antigens has been a major obstacle in the development of

* Corresponding author. Tel.: +1 617 876 8191x8203.
** Corresponding author. Tel.: +1 617 432 4156.
E-mail addresses: jessica.flechtner@genocea.com (J.B. Flechtner), darren_higgins@hms.harvard.edu (D.E. Higgins).

0264-410X/$ – see front matter © 2013 Elsevier Ltd. All rights reserved.
http://dx.doi.org/10.1016/j.vaccine.2013.06.046

T cell-based vaccines. T cell antigens are also potential targets for diagnosis of chronic infections and as biomarkers for understanding the pathogenesis of infectious and autoimmune diseases [7].

While Enzyme-Linked Immunosorbent Assays (ELISA) and protein microarrays can quickly identify antibody specificity, identifying T cell antigens relies on complex interactions between T cells and antigen presenting cells (APCs). Replicating these interactions in a laboratory has been a time-consuming and expensive process for several reasons. First, little is known about the characteristics of protective T cell antigens, so it is impossible to predict what might be a good antigen based on pathogen expression or localization. Second, due to Human Leukocyte Antigen (HLA) diversity, different individuals may respond to different portions of antigenic proteins. Thus, the conventional discovery of peptide epitopes may not be representative of the antigenic specificities of a broad population. Third, for large-genome organisms, the traditional approach of using overlapping peptides (OLPs) for comprehensive antigen discovery across the entire proteome is prohibitive due to both the cost and the number of patient cells required. Over the past decade, many different methods have been developed to identify T cell antigens in a high-throughput manner. In this review, we will cover some of these approaches, compare their advantages and disadvantages, and focus mainly on those methods designed to discover whole protein antigens that will have broad population coverage Table 1.

2. Traditional methods

The most common method for identifying T cell antigen specificities is based on the fact that T cell immunogens are derived from short amino acid sequences within each antigenic protein. Libraries of OLPs whose sequences span the entire protein of interest are synthesized and pulsed onto APCs, either individually or as pools. The ability of T cells to recognize OLPs is measured by the production of cytokines after co-culture with OLP-pulsed APCs. Two primary limitations of this approach are the restriction to specific HLA types used for epitope mapping, and the cost and time required to synthesize and screen large peptide libraries. Furthermore, generating libraries of OLPs for identifying T cell antigens for cancers or complex pathogens such as bacteria and parasites can be prohibitively expensive due to the larger size of the genome/proteome.

The cost of peptide libraries can be reduced, however, by limiting screening to selected sequences. Algorithms that predict the likelihood of a peptide to be bound by a major histocompatibility complex (MHC) molecule based on known specificities of MHC alleles are readily available to select those OLPs most likely to be recognized by T cells. These programs have been reviewed extensively elsewhere [8,9]. The algorithms take the peptide sequence of proteins and generate a list of likely antigens, greatly reducing the number of peptides that need to be tested. These algorithms still require a large number of peptides to be tested, far too many to explore the full proteomes of large-genome pathogens. In addition, MHC binding is only one of many factors that determine whether a particular peptide is antigenic and use of these algorithms may cause researchers to overlook antigens containing post-translational modifications such as phosphorylation. Consequently, predictive algorithms are limited by their inherent bias toward epitopes that may not represent the potential antigens across the entire proteome.

One of the first non-biased approaches used to identify T cell antigens examined epitopes directly from infected cells [10]. In this method, peptides are acid-eluted from the MHC complexes on the surface of pathogen-infected APCs. This pool of peptides is then fractionated by high-performance liquid chromatography (HPLC) and the fractions used to pulse additional APCs for a T cell restimulation assay. The antigenic peptide is identified from the activating fraction by mass spectrometry (MS). Since only a small amount of any given peptide can be recovered from the surface of APCs, biologically relevant antigens may not exist in sufficient quantities to restimulate T cells or be identified by MS. However, as advancements have been made in peptide detection by MS, many additional antigens are being discovered with this technique, including several promising vaccine candidates [11].

3. Whole-Protein methods

The goal of protein-based methods for T cell antigen discovery is the ability to rapidly and cost-effectively screen all, or a major fraction of the proteome of interest to identify key immunogens. All of the screens described below are focused on T cell responses. That is, the initial screen for antigens uses T cell stimulation, not MHC binding, as the readout for positive hits [7]. Generally, T cells from an exposed individual, either a human patient or an experimentally infected animal, are added to APCs capable of presenting a library of candidate antigens. If the natural exposure primed a T cell response against a particular candidate antigen, the APCs presenting that antigen will restimulate their cognate T cells *in vitro*. This restimulation is most commonly identified by measurement of secreted factors of activated T cells such as interferon (IFN)-γ or interleukin (IL)-17 using ELISA, Enzyme-Linked Immunosorbent Spot (ELISpot), or intracellular cytokine staining. This approach has the advantage of more closely mimicking natural T cell presentation than peptide-based pulsing, since the candidate proteins are delivered into the endogenous processing pathways of the APCs.

The major advantage of protein-based over peptide-based screens is that they are quicker and simpler. A single protein may contain many predicted peptide epitopes based on algorithms, but only a fraction or potentially none of them may be antigenic. Furthermore, subunit vaccines that are applicable across a broad population are often more dependent on the protein antigens than specific epitopes. Therefore, a protein-based screening approach can greatly simplify the number of targets a screen needs to encompass to adequately cover a pathogen's proteome. Traditionally, whole protein screens have used a variety of methods to limit the size of the library being screened. Many groups have focused $CD4^+$ T cell antigen screens on proteins that are also targets of the humoral immune response. This is done partially because of the principle of linked recognition, where an antigenic epitope for $CD4^+$ T cells must be physically associated with an antibody epitope to generate a response [12]. If T cells and antibodies are found to be directed to the same antigens, a subunit vaccine containing these antigens could prime a multifactorial response to the pathogen. For intracellular bacterial pathogens and eukaryotic parasites, $CD8^+$ T cell screens will often focus on proteins that are secreted by the pathogen. Such proteins are presumed to be more accessible to the APC's cytosolic MHC class I processing pathway than a protein retained within the pathogen's cytoplasm [13]. The end result is that while protein-based approaches can quickly identify T cell antigens, they still tend to result in some pre-selection or limited screening of predicted antigens.

One key concern for protein-based screens is the source of T cells used in the restimulation assay. If the screen is being performed using T cells isolated from animal models at the peak of infection, it may only be possible to screen the library using purified primary T cells. However, if the study is addressing a memory phenotype or using T cells isolated from patient blood samples, there may be insufficient quantities of primary T cells to screen the entire library. In this case, the T cells must be expanded. One common method for T cell expansion is to stimulate the T cells with immobilized anti-CD3 antibody and soluble cytokines. This method will expand all T

Table 1
Methods for high-throughput identification of T cell antigens.

Method	Vector	Antigen Presentation	Suitable for	Advantages	Disadvantages	Disease/Infection	References in this paper
Overlapping peptides	NA	Exogenous loading	Class I and II	No peptide mapping requirement	Slow, expensive, may not reflect natural antigen processing, limited HLA	Many	NA
Peptide prediction algorithms	NA	Exogenous loading	Class I and II	No peptide mapping requirement	Guesswork; peptide binding does not necessarily equate to protective immunity	Many	8,9
Peptide elution and mass spectrometry	NA	Pathogen-infected APC	Class I and II	Preserves natural antigen processing, directly identifies peptide antigens	May not identify less abundant antigens	Many	10,11
Purified protein	NA	APC uptake and endosomal processing	Class II only	Protein library can also be used for B cell antigen screening	Requires protein purification	*C. trachomatis*, Vaccinia, *A. marginale*	24–26
Bacteria expressing antigen	*E. coli*	APC uptake and endosomal processing	Class II only	Does not require protein purification	Random library requires deconvolution	*M. tuberculosis*, Vaccinia	27,28
Transfected cDNA library	Lentivirus	APC transfection and proteosomal processing	Class I only	cDNA library relatively simple to prepare	Difficult to gauge expression of antigens among APC clones	*C. trachomatis*, Vaccinia	20–22
ORFeome transfection	Lentivirus	APC transfection and proteosomal processing	Class I only	Ordered library does not require deconvolution	Not all potential antigens can be expressed in transfected APCs	Vaccinia	16
Synthetic minigene library	mRNA	Electroporation into B cells, proteosomal processing	Class I only	Well-suited to screen variants	Pooled antigens must be deconvoluted	Type 1 Diabetes	23
Expression Library Immunization (ELI)	DNA/VSV	Mammalian expression, APC uptake, endosomal or proteosomal processing	Class I and II	Identifies protective antigens	Requires extensive *in vivo* work, low throughput	Cowpox, *T. cruzi, M. bovis*, melanoma	41–46
T Cell Antigen Discovery (T-CAD)	NA	APC uptake and endosomal processing	Class I and II	Simple, colorimetric readout of T cell activation	Requires fusion of T cells to hybridomas and protein purification	*F. tularensis*, prostate cancer	17,30–32
Combined ORFeome transfection and purified protein screens	pDEST103	APC transfection and proteosomal or endosomal processing	Class I and II	Enhanced sensitivity using CD137 sorting	Shares disadvantages with Purified Protein and ORFeome Transfection methods; artificial APCs limit HLA types	HSV-1	15
AnTigen Lead Acquisition System (ATLAS™)	*E. coli*/cLLO	APC uptake, endosomal or proteosomal processing	Class I and II	Can use primary T cells as input; simple to screen for either class I or class II antigens, high throughput	Requires extensive cloning	HSV-2, *S. pneumoniae*, *C. trachomatis*,	36–40

cells non-specifically, preserving the antigen specificities present in the original isolate [14]. However, the expanded pool will include a large number of T cells that do not react to the pathogen of interest. An alternative is to incubate the T cells with APCs that have been pulsed with the particular pathogen [15]. The APCs should present the majority of the potential antigens from the pathogen and will stimulate pathogen-specific T cell expansion. However, the possibility still exists that the *in vitro* system does not accurately reflect *in vivo* antigen presentation.

While knowing which proteins are antigenic is the first step toward designing an effective subunit vaccine, diagnostic tools and further research often rely on knowing the specific antigenic determinant(s). To identify antigenic epitopes within an identified protein antigen, OLPs in combination with peptide prediction algorithms can be used [16]. Alternatively, some groups have subsequently used antigenic library expression clones to produce genetic C-terminal truncations of the antigens of interest [17,18]. These C-terminal truncations are then expressed in the same manner as the screening library and used to pulse APCs. A T cell line specific for the protein antigen of interest is added, and the T cell stimulation assay is repeated. The C-terminal truncations that elicit a T cell response contain the epitope of interest. The exact peptide epitope is then mapped using a small number of OLPs.

4. Methods suitable for MHC class I-restricted antigens

Since antigen processing for presentation by MHC class I molecules takes place in the cytosol of APCs, protein-based screens must first deliver any candidate antigens to this compartment for antigen processing to occur. One common way to accomplish this is to express candidate antigens recombinantly within the APCs, either by transformation with a viral expression vector or by transfecting the cell with nucleic acid encoding the protein of interest. However, not all candidate $CD8^+$ T cell antigens can be effectively expressed by mammalian cells. To overcome this limitation, other means, such as bacterial toxins, cross-presentation, and virus-like particles have been investigated to deliver protein expressed exogenously into the APC cytosol [19]. While most of these efforts have focused on vaccine delivery, some have been extended to antigen discovery.

Some of the first protein-based antigen screens used random libraries of pathogen cDNA [20] or genomic fragments [21] stably transfected into pools of APCs that were incubated with a clonal T cell line. Pools that activated the T cell line were used for further study. From these pools, antigens of interest were identified by performing limiting dilution of the APCs and sequencing clones that activated the T cell line. This method has allowed researchers to identify $CD8^+$ T cell antigens from *Chlamydia trachomatis* [20,22] and vaccinia virus [21]. However, the major drawback of using a random library is that the genomic coverage is unknown, and high-expressing antigens may be identified repeatedly at the expense of lower-expressing antigens. Furthermore, the random nature of the transfected pools requires that the library be large to ensure full coverage of the pathogen's potential antigenic sequences. Such large libraries require sizeable numbers of T cells and are extremely time-consuming to screen.

To resolve some of these issues, Bennick, Yewdell and colleagues have employed an ordered library of APCs each transfected with a lentiviral vector expressing a single protein from the pathogen of interest [16]. While an ordered library is more cumbersome and time-intensive to generate than other approaches described here, it can be more easily used to query mixed populations of T cells. Also, the specific proteins included in the expression library are known, allowing researchers to know the extent of the pathogen's proteome covered by the library. One disadvantage with this system is that it can be difficult to evaluate the level of expression of the candidate antigens between different APC clones.

Researchers in the Stone laboratory have developed a system for identifying $CD8^+$ T cell antigens using an ordered library of synthetic minigenes. The researchers chose 186 ORFs likely to encode autoimmune antigens and synthesized 33-codon DNA minigenes that covered all of the predicted proteins with 10 codon overlaps between minigenes. PCR was used to add a T7 promoter, Kozak sequence and polyadenyl sequence to each minigene, which were then pooled into groups of 10 and *in vitro* transcribed to produce mRNA. The mRNA pools were electroporated into B cells to use as APCs in a T cell activation assay. This method has the advantage of being able to quickly and relatively cheaply synthesize the library, saving time and effort over an ORF cloning approach. Due to the ease of synthesis, the library could also be used to examine a number of common variants of each gene included in the analysis, a distinct advantage for researchers looking for tumor- or autoimmunity-related antigens [23].

5. Methods suitable for MHC class II-restricted antigens

Since MHC class II processing takes place in the late endosome, whole proteins can be queried for their ability to stimulate T cells by simply adding them to phagocytic APCs. Most methods for identifying $CD4^+$ T cell antigens involve recombinant expression and purification of proteins to be added to APCs. A group from Novartis recently used this technique to test a panel of 79 *C. trachomatis* proteins purified from recombinant *Escherichia coli* (*E. coli*) for $CD4^+$ T cell antigens. From this panel, 21 proteins stimulated IFN-γ responses from $CD4^+$ T cells, of which 16 proved to be novel T cell antigens [24]. The disadvantage of this technique is that it requires the production and purification of each screened protein, which is expensive, time-consuming, and limits the number of proteins that can be evaluated.

Some research groups have had success combining recombinant protein screening with high-throughput protein production using *in vitro* transcription and translation (IVTT). This process can express a wider range of proteins than *E. coli* and reduces some of the cost burden, allowing for larger protein libraries. $CD4^+$ T cell epitopes from patients vaccinated against vaccinia virus [25] and cows vaccinated with preparations from *Anaplasma marginale* [26] have been discovered using IVTT-produced libraries of 200 and 50 different proteins, respectively.

As an alternative to purified proteins, libraries of *E. coli* clones containing expression vectors with genomic DNA fragments or cDNA constructs from the pathogen of interest have been used for screening. The *E. coli* clones are taken up by phagocytic APCs and lysed. Subsequently, bacterial contents enter the MHC class II processing pathway, and the APCs are screened for their ability to stimulate pathogen-specific T cells. This method has been used to identify antigens from *Mycobacterium tuberculosis* [27] and vaccinia virus [28]. Since this approach does not require protein purification, the library can be built and screened much more affordably than a protein library. One disadvantage of this technique is the reliance on random expression libraries, which makes deconvolution of antigenic specificities from a polyclonal T cell population difficult.

6. Methods suitable for both MHC class I and class II-restricted antigens

To identify both MHC class I and class II-restricted epitopes from vaccinia and Herpes simplex viruses, David Koelle and colleagues combined an ORFeome transfection method with purified protein screening. A library of pathogen ORFs was generated, which was

then used both to transfect APCs to discover CD8+ T cell antigens, and produce purified protein by IVTT to discover CD4+ T cell antigens [15]. This method also incorporated an additional strategy for discovering T cell antigens from infected patients. Instead of screening primary T lymphocytes, peripheral blood mononuclear cells (PBMC) were first stimulated *in vitro* with killed virus, then CD137+, CD8+ and CD4+ cells were purified and expanded using anti-CD3. Since CD137 is a marker of recently activated T cells [29], the expanded polyclonal pool has improved sensitivity when used in the restimulation assay. The CD137-enriched pool was able to identify T cell antigens that would have been missed had unsorted PBMCs been used. Rather than isolating APCs from peripheral blood, the combined approach uses Cos-7 cells transfected with HLA cDNA isolated from the patient being screened [15]. These artificial APCs avoid the need to differentiate APCs from each patient's peripheral blood sample and also allow examination of a single HLA allele at a time when screening the library, saving time on downstream characterization of each identified antigen.

Researchers in the Frelinger laboratory have identified several CD4+ and CD8+ T cell antigens from *Francisella tularensis* and prostate cancer [17,30,31] using a system called T-CAD (T-Cell Antigen Detection) [32]. This system uses a library of purified proteins conjugated to beads and added to APCs. The beads facilitate antigen uptake by APCs and increase antigen presentation through the MHC class II processing pathway. The bead conjugation also facilitates cross-presentation, where proteins from the endosome are loaded onto MHC class I molecules, allowing T-CAD to discover CD8+ as well as CD4+ T cell antigens. Instead of using primary or expanded T cells to screen the library, T cells from immunized animals are fused to the BWZ.36/CD8+ immortalized T cell line. BWZ.36/CD8+ cells contain a copy of the *E. coli lacZ* gene fused to a NFAT-dependent promoter. Thus, the resulting hybridomas will produce the enzyme β-galactosidase upon recognizing their cognate antigen. The level of β-galactosidase activity induced by each candidate antigen is then quickly measured with a colorimetric assay [33]. Screening with hybridomas has the advantage of generating robust, clonal T cell lines that can be used for further study of antigens. Since T-CAD can discover both CD4+ and CD8+ T cell antigens, any antigen identified must be assayed for presentation on MHC class $II^{-/-}$ APCs to determine MHC restriction [30].

Genocea Biosciences has developed a technology that uses an ordered proteomic library from the pathogen of interest, but does not require the purification of recombinant proteins. Rather, *E. coli* cells expressing the protein to be screened are fixed and then added directly to phagocytic APCs. The APCs lyse phagocytosed *E. coli* cells and process and present the antigens expressed within the bacteria. Pathogen-specific CD4+ T cells isolated from humans (or any species) can be added to the pulsed APCs and restimulation measured. To specifically identify CD8+ T cell antigens, the APCs are pulsed with an *E. coli* library co-expressing a modified version of the listeriolysin O (LLO) protein from *Listeria monocytogenes*. LLO is a pore-forming cytolysin with specificity for the phagosome [34,35]. The modified LLO protein used in the screen (cLLO) lacks a secretion signal and is confined to the *E. coli* cytoplasm. When the APC lyses the phagocytosed *E. coli*, cLLO is released along with the antigen of interest. cLLO perforates the phagosomal membrane and allows the antigen to enter the MHC class I processing pathway in the APC cytosol [36]. A major strength of this technology is that T cell antigen identification can be performed using diverse human subjects in the context of their natural exposure. There is no requirement for the generation of cell lines or for deconvolution of pools of antigens. This approach, termed ATLAS™ (AnTigen Lead Acquisition System), has been used successfully to discover CD8+ and CD4+ T cell antigens from *C. trachomatis* and Herpes Simplex Virus Type-2 [37,38] and CD4+ antigens from *Streptococcus pneumoniae* [39,40].

CD4+ and CD8+ T cell antigens from organisms including cowpox virus [41], *Trypansoma cruzi* [42], and *Mycobacterium bovis* [43] have been discovered using Expression Library Immunization (ELI). ELI has been reviewed in depth elsewhere [44,45], but the basic process is to generate a library of expression plasmids that encode the candidate proteins. This can be done using a cDNA library, sheared genomic DNA, or cloned ORFs. Plasmid pools are then used as DNA vaccines to immunize animals and then challenge with the pathogen of interest. Pools that induce protection are then deconvoluted to identify the most protective antigens. From a vaccine development standpoint, ELI has advantages in that it only identifies protective antigens that are amenable to vaccination. However, ELI has the disadvantage of requiring multiple expensive and time-consuming *in vivo* studies to identify the antigens, and for this reason, is not applicable to identification of antigens from humans. Vile and colleagues have updated the ELI approach to identify new CD4+ Th17 antigens from melanoma. The researchers isolated cDNA from a human melanoma cell line and used the cDNA to produce a vesicular stomatitis virus expression library. Mice injected with the B16 mouse melanoma line were then vaccinated with this library, curing the mice of melanoma. The researchers then isolated three viral clones capable of inducing IL-17 when used to infect cultures of splenocytes and lymph node cells obtained from mice previously injected with the B16 melanoma. While vaccination with any one of these clones individually failed to cure mice, combining the three clones into a single inoculum resulted in 60% of mice being cured of melanoma [46]. Thus, screening with viral clones allowed the researchers to find protective antigens that would not normally be discovered by ELI.

7. Conclusions

Whole protein T cell antigen screens have the potential to benefit basic research on host-pathogen interactions. The lists of T cell antigens that result from these screens can be used to generate tools such as MHC tetramers that can be used to probe the specific adaptive immune response to the pathogen. For some pathogens, namely vaccinia virus and *C. trachomatis*, many of the high-throughput approaches discussed in this review have been performed in multiple patients and animal models. This wealth of data, perhaps when combined with future experiments, may reveal underlying principles that define what proteins are effective T cell antigens. However, these principles may be pathogen-specific; certain proteins may not be available to the MHC processing pathways due to pathogen lifestyle or immune evasion mechanisms. Alternatively, T cell antigenicity may reflect regions of the protein sequence that are excluded from the T cell repertoire due to thymic selection. In either case, this information could be used to improve the next generation of antigen prediction software. If T cell antigens could be more accurately predicted using computer algorithms, the time investment and cost of antigen discovery would be reduced even further.

Most importantly, whole protein screening techniques are greatly increasing the pace of T cell antigen discovery for complex pathogens and cancer, and may ultimately bring the application of genome-based vaccinology approaches to the development of T cell vaccines to fruition. Already some of these techniques are being used to screen large numbers of patients with a single disease. These studies may be able to define T cell antigens that correlate with immunity or lighter disease burdens, and perhaps serve as the basis for new vaccines. Simple, inexpensive antigen discovery screens could also provide new opportunities in personalized medicine. Since the T cell antigens recognized by one person are very different than those recognized by another, low-cost high speed antigen screens could be used to characterize the T cell response of a

particular patient to a chronic disease such as cancer or autoimmunity. Immune therapies for these diseases could be tailored based on the specific antigen response.

References

[1] Fleischmann RD, Adams MD, White O, Clayton RA, Kirkness EF, Kerlavage AR, et al. Whole-genome random sequencing and assembly of *Haemophilus influenzae* Rd. Science 1995;269:496–512.

[2] Sette A, Rappuoli R. Reverse vaccinology: developing vaccines in the era of genomics. Immunity 2010;33:530–41.

[3] Davies DH, Liang X, Hernandez JE, Randall A, Hirst S, Mu Y, et al. Profiling the humoral immune response to infection by using proteome microarrays: high-throughput vaccine and diagnostic antigen discovery. Proceedings of the National Academy of Sciences of the United States of America 2005;102:547–52.

[4] Ottenhoff TH, Kaufmann SH. Vaccines against tuberculosis: where are we and where do we need to go? PLoS Pathogens 2012;8:e1002607.

[5] Mellman I, Coukos G, Dranoff G. Cancer immunotherapy comes of age. Nature 2011;480:480–9.

[6] Brunham RC, Rey-Ladino J. Immunology of Chlamydia infection: implications for a *Chlamydia trachomatis* vaccine. Nature Reviews Immunology 2005;5:149–61.

[7] Liu J, Zhang S, Tan S, Zheng B, Gao GF. Revival of the identification of cytotoxic T-lymphocyte epitopes for immunological diagnosis, therapy and vaccine development. Experimental Biology and Medicine 2011;236:253–67.

[8] Rosa DS, Ribeiro SP, Cunha-Neto ECD4+. T cell epitope discovery and rational vaccine design. Archivum Immunologiae et Therapiae Experimentalis 2010;58:121–30.

[9] Lundegaard C, Lund O, Buus S, Nielsen M. Major histocompatibility complex class I binding predictions as a tool in epitope discovery. Immunology 2010;130:309–18.

[10] Pamer EG. Direct sequence identification and kinetic analysis of an MHC class I-restricted *Listeria monocytogenes* CTL epitope. Journal of Immunology 1994;152:686–94.

[11] Karunakaran KP, Rey-Ladino J, Stoynov N, Berg K, Shen C, Jiang X, et al. Immunoproteomic discovery of novel T cell antigens from the obligate intracellular pathogen Chlamydia. Journal of Immunology 2008;180:2459–65.

[12] Parker DC. T cell-dependent B cell activation. Annual Review of Immunology 1993;11:331–60.

[13] Shen H, Miller JF, Fan X, Kolwyck D, Ahmed R, Harty JT. Compartmentalization of bacterial antigens: differential effects on priming of CD8 T cells and protective immunity. Cell 1998;92:535–45.

[14] Wang LX, Huang WX, Graor H, Cohen PA, Kim JA, Shu S, et al. Adoptive immunotherapy of cancer with polyclonal, 108-fold hyperexpanded, CD4+ and CD8+ T cells. Journal of Translational Medicine 2004;2:41.

[15] Jing L, Haas J, Chong TM, Bruckner JJ, Dann GC, Dong L, et al. Cross-presentation and genome-wide screening reveal candidate T cells antigens for a herpes simplex virus type 1 vaccine. The Journal of Clinical Investigation 2012;122:654–73.

[16] Tscharke DC, Karupiah G, Zhou J, Palmore T, Irvine KR, Haeryfar SM, et al. Identification of poxvirus CD8+ T cell determinants to enable rational design and characterization of smallpox vaccines. The Journal of Experimental Medicine 2005;201:95–104.

[17] Valentino MD, Abdul-Alim CS, Maben ZJ, Skrombolas D, Hensley LL, Kawula TH, et al. A broadly applicable approach to T cell epitope identification: application to improving tumor associated epitopes and identifying epitopes in complex pathogens. Journal of Immunological Methods 2011;373:111–26.

[18] Jing L, Schiffer JT, Chong TM, Bruckner JJ, Davies DH, Felgner PL, et al. CD4 T-cell memory to viral infections of humans shows pronounced immunodominance independent of duration or viral persistence. Journal of Virology 2012.

[19] Moron G, Dadaglio G, Leclerc C. New tools for antigen delivery to the MHC class I pathway. Trends in Immunology 2004;25:92–7.

[20] Fling SP, Sutherland RA, Steele LN, Hess B, D'Orazio SE, Maisonneuve J, et al. CD8+ T cells recognize an inclusion membrane-associated protein from the vacuolar pathogen Chlamydia trachomatis. Proceedings of the National Academy of Sciences of the United States of America 2001;98:1160–5.

[21] Jing L, Chong TM, McClurkan CL, Huang J, Story BT, Koelle DM. Diversity in the acute CD8 T cell response to vaccinia virus in humans. Journal of Immunology 2005;175:7550–9.

[22] Starnbach MN, Loomis WP, Ovendale P, Regan D, Hess B, Alderson MR, et al. An inclusion membrane protein from *Chlamydia trachomatis* enters the MHC class I pathway and stimulates a CD8+ T cell response. Journal of Immunology 2003;171:4742–9.

[23] Hondowicz BD, Schwedhelm KV, Kas A, Tasch MA, Rawlings C, Ramchurren N, et al. Discovery of T cell antigens by high-throughput screening of synthetic minigene libraries. PloS ONE 2012;7:e29949.

[24] Finco O, Frigimelica E, Buricchi F, Petracca R, Galli G, Faenzi E, et al. Approach to discover T- and B-cell antigens of intracellular pathogens applied to the design of *Chlamydia trachomatis* vaccines. Proceedings of the National Academy of Sciences of the United States of America 2011;108:9969–74.

[25] Jing L, McCaughey SM, Davies DH, Chong TM, Felgner PL, De Rosa SC, et al. ORFeome approach to the clonal, HLA allele-specific CD4 T-cell response to a complex pathogen in humans. Journal of Immunological Methods 2009;347:36–45.

[26] Lopez JE, Beare PA, Heinzen RA, Norimine J, Lahmers KK, Palmer GH, et al. High-throughput identification of T-lymphocyte antigens from *Anaplasma marginale* expressed using in vitro transcription and translation. Journal of Immunological Methods 2008;332:129–41.

[27] Coler RN, Dillon DC, Skeiky YA, Kahn M, Orme IM, Lobet Y, et al. Identification of Mycobacterium tuberculosis vaccine candidates using human CD4+ T-cells expression cloning. Vaccine 2009;27:223–33.

[28] Jing L, Chong TM, Byrd B, McClurkan CL, Huang J, Story BT, et al. Dominance and diversity in the primary human CD4 T cell response to replication-competent vaccinia virus. Journal of Immunology 2007;178:6374–86.

[29] Wehler TC, Karg M, Distler E, Konur A, Nonn M, Meyer RG, et al. Rapid identification and sorting of viable virus-reactive CD4(+) and CD8(+) T cells based on antigen-triggered CD137 expression. Journal of Immunological Methods 2008;339:23–37.

[30] Valentino MD, Hensley LL, Skrombolas D, McPherson PL, Woolard MD, Kawula TH, et al. Identification of a dominant CD4 T cell epitope in the membrane lipoprotein Tul4 from Francisella tularensis LVS. Molecular Immunology 2009;46:1830–8.

[31] Valentino MD, Maben ZJ, Hensley LL, Woolard MD, Kawula TH, Frelinger JA, et al. Identification of T-cell epitopes in *Francisella tularensis* using an ordered protein array of serological targets. Immunology 2011;132:348–60.

[32] Turner MJ, Abdul-Alim CS, Willis RA, Fisher TL, Lord EM, Frelinger JG. T-cell antigen discovery (T-CAD) assay: a novel technique for identifying T cell epitopes. Journal of Immunological Methods 2001;256:107–19.

[33] Sanderson S, Shastri N. LacZ inducible, antigen/MHC-specific T cell hybrids. International Immunology 1994;6:369–76.

[34] Beauregard KE, Lee KD, Collier RJ, Swanson JA. pH-dependent perforation of macrophage phagosomes by listeriolysin O from *Listeria monocytogenes*. The Journal of Experimental Medicine 1997;186:1159–63.

[35] Portnoy DA, Jacks PS, Hinrichs DJ. Role of hemolysin for the intracellular growth of *Listeria monocytogenes*. The Journal of Experimental Medicine 1988;167:1459–71.

[36] Higgins DE, Shastri N, Portnoy DA. Delivery of protein to the cytosol of macrophages using *Escherichia coli* K-12. Molecular Microbiology 1999;31:1631–41.

[37] Picard MD, Cohane KP, Gierahn TM, Higgins DE, Flechtner JB. High-throughput proteomic screening identifies *Chlamydia trachomatis* antigens that are capable of eliciting T cell and antibody responses that provide protection against vaginal challenge. Vaccine 2012;30:4387–93.

[38] Skoberne M, Cardin R, Lee A, Kazimirova A, Zielinski V, Garvie D, et al. An adjuvanted herpes simplex virus type 2 (HSV-2) subunit vaccine elicits a T cell response in mice and is an effective therapeutic vaccine in guinea pigs. Journal of Virology 2013.

[39] Moffitt KL, Gierahn TM, Lu YJ, Gouveia P, Alderson M, Flechtner JB, et al. T(H)17-based vaccine design for prevention of *Streptococcus pneumoniae* colonization. Cell host & Microbe 2011;9:158–65.

[40] Li Y, Gierahn T, Thompson CM, Trzcinski K, Ford CB, Croucher N, et al. Distinct effects on diversifying selection by two mechanisms of immunity against *Streptococcus pneumoniae*. PLoS Pathogens 2012;8:e1002989.

[41] Borovkov A, Magee DM, Loskutov A, Cano JA, Selinsky C, Zsemlye J, et al. New classes of orthopoxvirus vaccine candidates by functionally screening a synthetic library for protective antigens. Virology 2009;395:97–113.

[42] Tekiel V, Alba-Soto CD, Gonzalez Cappa SM, Postan M, Sanchez DO. Identification of novel vaccine candidates for Chagas' disease by immunization with sequential fractions of a trypomastigote cDNA expression library. Vaccine 2009;27:1323–32.

[43] Beltran PK, Gutierrez-Ortega A, Puebla-Perez AM, Gutierrez-Pabello JA, Flores-Valdez MA, Hernandez-Gutierrez R, et al. Identification of immunodominant antigens of *Mycobacterium bovis* by expression library immunization. Veterinary Journal 2011;190:181–3.

[44] Barry MA, Howell DP, Andersson HA, Chen JL, Singh RA. Expression library immunization to discover and improve vaccine antigens. Immunological Reviews 2004;199:68–83.

[45] Talaat AM, Stemke-Hale K. Expression library immunization: a road map for discovery of vaccines against infectious diseases. Infection and Immunity 2005;73:7089–98.

[46] Pulido J, Kottke T, Thompson J, Galivo F, Wongthida P, Diaz RM, et al. Using virally expressed melanoma cDNA libraries to identify tumor-associated antigens that cure melanoma. Nature Biotechnology 2012;30:337–43.