FEBS Letters 588 (2014) 318–325

journal homepage: www.FEBSLetters.org

Review

Profiling the IgOme: Meeting the challenge ☆

Yael Weiss-Ottolenghi, Jonathan M. Gershoni *

Department of Cell Research and Immunology, George S. Wise Faculty of Life Sciences, Tel Aviv University, Tel Aviv 69978, Israel

ARTICLE INFO

Article history:
Received 30 September 2013
Revised 6 November 2013
Accepted 6 November 2013
Available online 12 November 2013

Edited by Wilhelm Just

Keywords:
Antibody
Serum
phage display
Next generation sequencing
Serology
B-cell
Immunoglobulin

ABSTRACT

The entire repertoire of antibodies in our serum, the IgOme, is a historical record of our past experiences and a reflection of our immune status at any given moment. Understanding the dynamics of the IgOme and how the diversity and specificities of serum antibodies change in response to disease and maintenance of homeostasis can directly impact the ability to design and develop novel vaccines, diagnostics and therapeutics. Here we review both direct and indirect methodologies that are being developed to map the complexity and specificities of the antibodies in polyclonal serum – the IgOme.

1. Introduction

Over a Century has passed since the first Nobel Prize in Medicine (1901) was awarded to Emil von Behring for "his work on serum therapy ... by which he opened a new road in the domain of medical science and thereby placed in the hands of the physician a victorious weapon against illness and death" [1]. Serum is indeed a "victorious weapon" able to neutralize pathogens through the exquisite specificity of its antibodies that seem endless in their capacity to discriminate and bind the vast structural complexities found in nature. Susumu Tonegawa's Nobel Prize (1987) recognized the elucidation of the genetic principles for the generation of this antibody diversity [2]. Application of these principles allows the derivation of the total "potential repertoire" of antibodies in humans; the combinatorial multiplication of all the V, D, and J segments of the Heavy (H) chains times the combined product of the V and J segments of the kappa and lambda Light (L) chains as well as the contributions of the N and P nucleotides associated with the junctional complexity – leading to an extraordinary vast theoretical number, 10^{11}–10^{12}, that far exceeds the total number of B-cells in a person's body [34,65]. More realistic is the "available repertoire" of variant B-cells that has been calculated to be in the order of 10^7 per person [9]. How many of these B-cells subtypes are actually utilized for the production of distinct antibodies probably does not exceed more than tens of thousands in a person's life time. Surprisingly, whereas pathogens come and go as they are cleared from our bodies, the antibodies generated in response to immunological insults are archived in our memory B-cells, the cells that orchestrate the continuous production of antibodies found in serum over the course of our lives. This entire utilized repertoire of antibodies in our serum, the IgOme, is a historical record of our past experiences and a reflection of our immune status at any given moment.

Understanding the IgOme, how homeostasis is maintained, how "serum memory" is affected by immunization, boosts, encounters with pathogens, physiology and old age are all fundamental questions of great interest. The answers to these questions bear directly on the development of novel vaccines, diagnostics and therapeutics and in order to meet these challenges one must be able to profile the IgOme. We need to be able to describe the IgOme in its entirety at single antibody resolution in a manner that is cost effective and expedient. One could then imagine running IgOme screens routinely in the course of personalized medical diagnosis and treatment. Obviously, the challenge is formidable – just considering the diversity of antibodies and the dynamics of their expression at any given moment and as a result to any given stimulus. Here we review a variety of technical approaches designed to profile serum antibody diversity in the quest of describing the complexity and composition of the human IgOme.

☆ This study was supported by the generous support of the Frankel Foundation. Jonathan M. Gershoni is the incumbent of the David Furman Chair of Immunobiology in Cancer.

* Corresponding author. Fax: +972 3 642 2046.
E-mail addresses: gershoni@post.tau.ac.il, gershoni@tauex.tau.ac.il (J.M. Gershoni).

0014-5793/$36.00
http://dx.doi.org/10.1016/j.febslet.2013.11.005

2. Direct and indirect analysis of the IgOme

Two basically different approaches have been employed in the analysis of the antibody repertoire of serum. The "Direct" approach goes to profile the antibodies or antibody-producing cells themselves. For this, one can catalogue antibodies based on proteomic sequencing or achieve the same by genomic sequencing of antibody transcripts derived from peripheral B-cells. The ultimate goal would be to rank the antibodies by relative concentration and categorize them by antigen specificity. The "Indirect" approach uses surrogate measures to extrapolate antibody species based on proteins or peptides they bind. Hence, the IgOme profile would be a list of inferred antibody specificities. First we address the "Direct" analysis of the IgOme.

3. Direct IgOme profiling

3.1. Phage display of antibodies

Conceptually, profiling the IgOme in its entirety would be complete were one able to systematically clone the antibodies in polyclonal serum, one by one and catalogue them by frequency of usage along with their corresponding antigen specificities. Clearly, classical methods for the production of monoclonal Antibodies (mAbs) are not suitable for such massive comprehensive screening, not in mice and definitely not for humans (see for example reviews [40,58,97]).

The introduction of phage display libraries of antibodies, however, provided a breakthrough in that one could, in theory, preserve the entire repertoire of the Heavy (H) and Light (L) chains of a person to be later screened for various antigen specificities [16,43a,60]. Two types of antibody phage display libraries can be constructed: "naïve" and "immunized" libraries. Both types of libraries are generated to discover details of the immune response to a given pathogen, autoimmune disease or cancer [5,20,21,26,46,56,86,102]. Naïve libraries generated from random donors are presumed to represent the "healthy" human repertoire of antibodies in general. In principle, naïve libraries offer the possibility of selection of high affinity antibodies of any desired specificity without the need for selectively stimulating the immune response with a specific antigen. Using naïve libraries, there has been considerable success for isolating mAbs against various pathogens, such as SARS-CoV [69,86], WNV [37], and HBV [49]. The second class of library, the "immunized phage-displayed antibody library" is constructed from the mRNA from a donor who was exposed to infection/vaccination with a defined antigen, therefore his immune response has been specifically stimulated. For immunized libraries, it is generally preferable to use a donor with a high serum antibody titer for the antigen/pathogen of interest. A high serum titer is presumed to reflect relatively high levels of Ab production and therefore higher levels of specific mRNA should be obtainable for the generation of the library [17]. This approach has been successful for isolating mAbs specific for diverse pathogens, such as, H5N1 [46,90], foot and mouth disease [31] and HIV [6,23,41,64,101,102].

In the matter of HIV-1, isolation of potent neutralizing mAbs has only been successful using immunized libraries (as compared to naïve libraries). These have been constructed from selected donors that have proven neutralizing activity. A case in point is the study by Burton and Barbas where an antibody library was produced from a 31-year old, HIV-1 positive, homosexual male, who had been asymptomatic for six years [5,16]. A Fab library was constructed on the surface of filamentous phage which comprised 10^7 members. This library was then screened against monomeric glycosylated HIV-1 $gp120_{IIIB}$ from which a collection of 20 phage displayed mAbs was isolated. The most potent was mAb b12, which competes for the binding site of HIV-1 gp120 receptor (CD4) and thus prevents virus binding to its target cell. Co-crystallization of this mAb with core gp120 confirmed the epitope overlap with the CD4 binding site [105]. Moreover, these b12 studies were revealing of two major drawbacks of the phage display technology: (i) This methodology is still relatively time consuming, tedious, and somewhat inefficient. Typically in standard experiments only tens of mAbs are isolated which are specific towards the antigen against which they were screened. (ii) The natural pairing of immunoglobulin H and L chains is lost in the construction of the library, thus the vast majority of the antibodies produced are the result of random, arbitrary pairing which does not reflect the natural H:L pair of the B-cell clone that produced them. Indeed, the co-crystallization of mAb b12 shows that the L chain does not contribute to gp120 recognition at all. All the contacts are made exclusively *via* the H chain [105]. Thus, indicating that the L chain used is most probably not the H chain's natural partner. Generally, it has been estimated that when constructing an Ab phage display library from total H and L chain cDNAs derived from Peripheral Blood Mononuclear Cells (PBMCs) and generating a complexity of 10^8 phage displayed antibodies, only 10 000 mAbs are expected to maintain their native H:L pair [40].

Another drawback of the phage display antibody system is that some VH:VL pairs may be toxic and impair bacterial growth. This in turn may lead to a very biased representation of antibodies when phage display is used for antibody expression [75].

Nonetheless, the power of phage-displayed antibody libraries is the ability to immortalize the entirety of the available naïve potential of H and L chains and even expand upon it by generating novel pairs that do not naturally exist in the donor's repertoire.

3.2. Next Generation Sequencing of antibody H and L chain mRNAs

The premise for these analyses is that one can apply Next Generation Sequencing (NGS) to catalogue the entire complexity of antibody related transcripts of a given individual [13,34,96]. For this, high throughput sequencing is typically performed using the 454 pyrosequencing analyses that currently generate $>10^6$ sequences of 400–500 bp long. Thus, one can read beyond the three CDRs of a chain in a single run [3,8,13,34,96,100,106].

When total PBMCs are used, such analyses typically address the available repertoire and not necessarily the utilized repertoire. Thus for example, Glanville et al. conducted their analyses on pooled phage displayed IgM antibodies derived from 654 healthy human donors [34]. This study serves as a comprehensive and detailed analysis of the "available IgOme" and produces a baseline of sorts for the "average" human repertoire. The total diversity of the combinatorial Ab library was calculated to be 3.5×10^{10}. This library, when screened using 16 different antigens, produced a collection of antigen specific antibodies that when analyzed using NGS, were found to comprise all the human germline VDJ and VJ segments as expected. Moreover, this analysis provides insights to the relative frequency of the usage of V segments.

Application of NGS to the study of antibodies has increased our understanding of the general physical traits of antibodies produced in humans, such as: prevalence of usage of various germline V, D and J segments, for a given class of antigen [100]. Another subject of interest has been the specific focus on the characterization of HCDR3 loops, regarding their amino acid lengths and the number of "in-dels" detected in the generation of extended CDR3 loops [3,14,20,34,100]. Moreover, it emphasizes the power of NGS technology that provides a quality assurance to identify and correct biases that may be introduced during the procedure of library construction such as over representation or the absence of a specific segment [34].