Mining the immune system for protective antibodies

Genomics England
4 May 2022
OUR VISION: TO USE THE POWER OF THE HUMAN IMMUNE SYSTEM TO DISCOVER NEW MEDICINES AND DIAGNOSTICS

Discovering and developing protective, patient-originated therapeutic antibodies

- Focus on protective antibody responses
- Convergent in resilient individuals
- Target agnostic approach
- Advanced computational approaches
B CELL RECEPTORS ARE A KEY COMPONENT OF ADAPTIVE IMMUNITY

<table>
<thead>
<tr>
<th>INNATE</th>
<th>NONSPECIFIC</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fast response (0-4 hours)</td>
<td></td>
</tr>
</tbody>
</table>

- **MONONUCLEAR PHAGOCYTE SYSTEM**
  - Macrophage
  - Dendritic Cell
  - Monocyte
  - Complement Protein

- **Natural Killer Cell**
- **Mast Cell**
- **Basophil**
- **Eosinophil**
- **Neutrophil**

- **granulocites**

<table>
<thead>
<tr>
<th>ADAPTIVE</th>
<th>SPECIFIC</th>
</tr>
</thead>
<tbody>
<tr>
<td>Slow response (4-14 days)</td>
<td></td>
</tr>
</tbody>
</table>

- **HUMORAL**
  - **B cell**
  - **Antibodies**

- **CELLULAR**
  - **T Lymphocyte**
  - **CD4+**
  - **CD8+**

- **ϒδ T Cell**
- **Natural Killer T Cell**
“The sum of all B cells expressing distinct B cell receptors, necessary to bind diverse antigens and produce an effective humoral immune response”
THE B CELL RECEPTOR REPERTOIRE HAD TWO ELEMENTS

- Plasma Cells & Memory cells
- Secondary Repertoire (antigen-dependent)
- Primary Repertoire (antigen-independent)
- Stem cells
- Antigen
- Plasma Cells & Memory cells
FORMING THE PRIMARY REPERTOIRE

Primary Repertoire (antigen-independent)

Secondary Repertoire (antigen-dependent)

Stem cells

1

2

3

4

5

6

1

2

3

4

5

6

Antigen

Plasma Cells
&
Memory cells

(antigen-dependent)

(antigen-independent)
COMBINATORIAL DIVERSITY – VDJ RECOMBINATION

Somatic recombination of gene segments during B cell development

<table>
<thead>
<tr>
<th>Element</th>
<th>H</th>
<th>K+ λ</th>
</tr>
</thead>
<tbody>
<tr>
<td>Variable segments (V)</td>
<td>~40</td>
<td>~70</td>
</tr>
<tr>
<td>Diversity segment (D)</td>
<td>23</td>
<td>0</td>
</tr>
<tr>
<td>D segments read in three frames</td>
<td>rarely</td>
<td>-</td>
</tr>
<tr>
<td>Joining segments (J)</td>
<td>6</td>
<td>5(k) 4(λ)</td>
</tr>
<tr>
<td>Joints with N- and P-nucleotides</td>
<td>2</td>
<td>50% of joints</td>
</tr>
<tr>
<td>Number of V gene pairs</td>
<td>$1.9 \times 10^6$</td>
<td></td>
</tr>
<tr>
<td>Junctional diversity</td>
<td>~$3 \times 10^7$</td>
<td></td>
</tr>
<tr>
<td>Total diversity</td>
<td>~$5 \times 10^{13}$</td>
<td></td>
</tr>
</tbody>
</table>
FORMING THE SECONDARY REPERTOIRE

Primary Repertoire (antigen-independent)

Secondary Repertoire (antigen-dependent)

Antigen

Stem cells

Plasma Cells & Memory cells
AFFINITY MATURATION IN THE GERMINAL CENTRE

RAPID B CELL EVOLUTION OVER THE COURSE OF DAYS

IgM is the first antibody to appear in the primary immune response.

Class-switched plasma Cells and memory cells

First exposure to antigen
BCR HIGH THROUGHPUT SEQUENCING: HAS BEEN EVOLVING SINCE 2009

High-Throughput Sequencing of the Zebrafish Antibody Repertoire

Joshua A. Weinstein, 1,* Ming Jiang, 1,2,4,5 Richard A. White III, 1
Daniel S. Fisher, 2,4,5 Stephen R. Quake 1,2,4,5

Despite tremendous progress in understanding the nature of the immune system, the full diversity of an organism’s antibody repertoire is unknown. We used high-throughput sequencing of the variable domain of the antibody heavy chain from 15 zebrafish to analyze VDJ usage and antibody sequence. Zebrafish were found to use between 50 and 86% of all possible VDJ combinations and shared a similar frequency distribution, with some correlation of VDJ patterns between individuals. Zebrafish antibodies retained a few thousand unique heavy chains that also exhibited a shared frequency distribution. We found evidence of convergence, to a certain extent, in different individuals made the same antibody. This approach provides insight into the breadth of the expressed antibody repertoire and immunological diversity at the level of an individual organism.

The nature of the immune system’s antibody repertoire has been a subject of fascination for more than a century. This repertoire is highly plastic and can be directed to create antibodies with broad chemical diversity and high specificity (1, 2). There is also a good understanding of the potential diversity available and the mechanistic aspects of how this diversity is generated. Antibodies are composed of two types of chains (heavy and light), each containing a highly diversified antigen binding domain (V), the V, D, and J gene segments of the antibody heavy-chain variable genes go through a series of recombination events to generate a new heavy-chain gene (Fig. 1). Antibodies are formed by a mixture of recombination among gene segments, sequence diversification at the junctions of these segments, and point mutations throughout the gene (3). Estimates of immune diversity for antibodies or the related T cell repertoire either have attempted to extrapolate from small samples to entire species or have been limited by coarse resolution of immune repertoire genes (4). However, certain very elementary questions have remained open since the half-century after being posed (1, 5). It is still unclear what fraction of the potential repertoire is expressed in an individual at any point in time and how similar repertoires are between individuals who have lived in similar environments. Moreover, because each individual’s immune system is an independent experiment in evolution by natural selection, these questions about repertoire similarity also inform our understanding of evolutionary diversity and convergence.

Zebrafish are an ideal model system for studying the adaptive immune system because in evolutionary terms they have the earliest mappable adaptive immune system whose features match the essential human elements (1, 9). Like humans, zebrafish have a somatic recombination activating gene (RAG) and a combinatorial rearrangement of V, D, and J gene segments to create antibodies. They also have functional diversity during recombination and somatic hypermutation of antibodies to improve specificity, and the organization of their immunoglobulin (Ig) gene loci approximates that of mammals (9). In addition, the zebrafish immune system has only ~30,000 antibody-producing B cells, making it three orders of magnitude smaller than mouse and five orders smaller than human in this regard.

We developed an approach to characterize the antibody repertoire of zebrafish by analyzing comprehensive sequencing region 3 (CDR3) of the heavy chain, which contains the vast majority of immunoglobulin diversity (10). I can be captured in a single sequencing read (Fig. 1). Using the 454 FLX high-throughput sequencing technology allowed sequencing of 480 million bases of zebrafish antibody cDNA from 15 zebrafish in four families (Fig. 1B). Zebrafish were raised in separate aquaria for each family and were allowed to have normal interactions with the environment, including the development of natural immune flora. We chose to investigate the quiescent state of the immune system, a state where the zebrafish had sampled a complex but fairly low-immune environment and had established an equilibrium of normal immune function. DNA was prepared from whole fish, and we synthesized cDNA using primers designed to capture the entire sample cohort.

Between 20,000 and 12,000 useful sequencing reads were obtained per fish, and we focused our analysis on CDR3 segments. Each read was aligned to a representative reference with a 99.4% success rate (table S3). Failures were due to similarity in some of the V gene segments. D was determined for each read by applying a clustering algorithm to all of the reads within a given
HOW RARE IS CONVERGENCE?

DIVERSITY OF THE HUMAN REPERTOIRE IS THEORETICALLY $1 \times 10^{13}$

- Number of peripheral B cells in a healthy individual is approx. $1 \times 10^9$
- Circulating B Cell Repertoire is therefore a fraction of the total diversity
- The amount of information encoded in “genome” of the adaptive immune system exceeds the human genome by 4 orders of magnitude

10 INDIVIDUALS VH SEQUENCED
3 BILLION HEAVY CHAINs
Briney et al Nature 2019

- Largely unique repertoires for each individual
- Between 2 individuals 0.95% of the repertoire was shared. Shared clonotypes were skewed towards short CDR3s
- Only 0.022% of clones were shared between all 10 individuals
- Commonality is driven by early BCR development rather than common antigen-specific selection, although there is some convergence due to vaccination and common infections
Convergence increases after a common antigen stimulus and can be used to identify rare antigen-specific sequences.

Vaccination used as an example of a strong, specific stimulus.
AUTOANTIBODIES ARE GENERALLY VIEWED AS ‘BAD GUYS’

MANY AUTOIMMUNE DISEASES ARE THE RESULT OF PATHOGENIC AUTOANTIBODIES:

- SLE (Polyclonal autoreactive B cells)
- Myasthenia Gravis (anti-AchR)
- Neurological disease (anti-NMDA Receptor encephalitis)
- Pemphigus (anti-desmosomal proteins)

TREATMENTS:

- Ablate B cells (anti-CD20)
- Plasmapheresis
- Anti-FcRn
- High dose IVIG to block
- Anti-IDs (e.g. anti-anti-Dsg3 autoantibodies)
IF THERE ARE BAD GUYS THERE ARE ALWAYS GOOD GUYS

Published: 19 March 2013

Protective autoantibodies in the rheumatic diseases: lessons for therapy

Gregg J. Silverman, Jaya Vas & Caroline Grönwall

Nature Reviews Rheumatology 9, 291–300 (2013) | Cite this article
1639 Accesses | 32 Citations | 1 Altmetric | Metrics

ALDH4A1 is an atherosclerosis auto-antigen targeted by protective antibodies

https://doi.org/10.1038/nrheum.2020.239
Received: 3 December 2019
Accepted: 5 October 2020
Published online: 2 December 2020

Protective effect of naturally occurring anti-HER2 autoantibodies on breast cancer

Yukiko Tabuchi¹ · Masafumi Shimoda¹ · Naofumi Kagara¹ · Yasuto Naoi¹ · Tomonori Tanii³ · Atsushi Shimomura¹ · Kenzo Shimazu¹ · Seung Jin Kim¹ · Shinzaburo Noguchi²

DOI 10.1007/s10519-016-3801-4

Recurrence-free survival (%)
ALCHEMAB CONCEPT

WE IDENTIFY
We identify especially resilient individuals – and learn how they overcome or resist disease

WE SEQUENCE
We sequence B cells from the resilient individuals and identify antibodies with similar properties

WE DISCOVER
We discover the binding targets of the antibodies, understand their protective properties and develop candidates that replicate the protective effect

Unbiased platform to identify novel therapeutics
DEFINING “RESILIENCE”

Patients who survive advanced cancer

Patients progressing unusually slowly with neuro-degenerative disorders

People who survive grievous, deadly infectious disease

Long-lived, healthy individuals
WE COLLABORATE WITH A BROAD AND GROWING NETWORK OF INSTITUTIONS

NEURODEGENERATION

CHDI Foundation

EPAD

Queen Mary University of London

UCL

PLATFROM TECHNOLOGY

ONCOLOGY

HCL Hospices Civils de Lyon

ST MARKS

Cancer Research UK

Gustave Roussy

AstraZeneca

OTHER AREAS

Johns Hopkins University

illuminA

University of Oxford

Amsterdam UMC

NHS Barts Health

NVIDIA
CONVERGENT AUTOANTIBODY SEQUENCES PROVIDE THE STARTING POINT FOR DISCOVERY

POPULATION B CELL REPERTOIRE ANALYSIS

Sequences of disease-free population
Sequences of resilient patients
Sequences of disease progressors

10's-100's of Individuals; Millions of BCR sequences

CONVERGENCE

Shared antibodies that do not occur in controls or progressors
REDUCING THE VAST B CELL REPERTOIRE DOWN TO CANDIDATE ANTIBODIES

CLUSTERING
Grouping together antibodies based on sequence relatedness

CONVERGENCE
Searching for clusters containing sequences derived from multiple resilient individuals – indicates selection for similar specificities

VALIDATE & TRIAGE
Validating convergent signals, and understanding features of the best antibodies
DISCOVERY PROCESS POWERED BY ADVANCED SEQUENCING, BIG DATA, AND DEEP LEARNING

CORE TECHNOLOGIES

1. Repertoire analysis
   - Leading expertise & capabilities in BCR repertoire analysis

2. Antigen proteomics
   - Serum analysis and target deconvolution tools

3. Deep learning
   - Proprietary technologies

PROCESS

Samples

Data driven phenotypic and functional biology cascade
DISEASE-ASSOCIATED ANTIBODY SIGNATURES

Shared by individuals within a disease group and distinct from healthy controls.

Heatmap of 1,337 COVID-19 clonotypes, 1,180 Influenza clonotypes and 351 breast cancer clonotypes, demonstrating that the convergent signatures are unique to each disease cohort. Healthy controls subtracted.

*Breast cancer
Infectious Disease
CASE STUDY
Potent multi-strain covid neutralizing antibodies identified

Deep Sequencing of B Cell Receptor Repertoires From COVID-19 Patients Reveals Strong Convergent Immune Signatures

Jacob D. Galson¹,², Sebastian Schaeztle³, Rachael J. M. Bashford-Rogers¹,², Matthew I. J. Raybould⁴, Aleksandr Kovaltsuk⁵, Gavin J. Kilpatrick⁶, Ralph Minter⁷, Donna K. Finch⁸, Jorge Dias⁹, Louisa K. James⁹, Gavin Thomas⁹, Wing-Yiu Jason Lee³, Jason Betley³, Olivia Cawlan¹, Alex Leech¹, Charlotte M. Deane⁶, Joan Seoane⁶, Carlos Caldas¹, Daniel J. Pennington⁴, Paul Pfeffer⁴ and Jane Osbourn¹

Galson et al., https://doi.org/10.3389/fimmu.2020.605170
Oncology
CASE STUDY
ALL PDAC PATIENTS HAVE UNDERGONE RESECTION
HIGHLY CONVERGENT RESILIENT CLONOTYPES FOUND IN MULTIPLE INDIVIDUALS

10 clonotypes found with exceptional convergence among resilient individuals which were not found in progressors:

- Length-independent super-convergence
- Low probability of generation & rare in healthy controls
- Predominantly IgG1

<table>
<thead>
<tr>
<th>CLUSTER SIZE</th>
<th>CONVERGENCE LEVEL (OUT OF 12)</th>
<th>CDR3 LENGTH</th>
<th>MUTATIONS</th>
<th>GENERATION PROBABILITY</th>
<th>PROP. IN HEALTHY CONTROL</th>
<th>PREVALENT ISOTYPE</th>
</tr>
</thead>
<tbody>
<tr>
<td>151</td>
<td>8</td>
<td>20</td>
<td>1.58</td>
<td>4.35E-12</td>
<td>0.060</td>
<td>IGHG1</td>
</tr>
<tr>
<td>40</td>
<td>8</td>
<td>20*</td>
<td>1.65</td>
<td>9.34E-12</td>
<td>0.050</td>
<td>IGHG1</td>
</tr>
<tr>
<td>43</td>
<td>8</td>
<td>19*</td>
<td>1.35</td>
<td>4.63E-20</td>
<td>0.186</td>
<td>IGHG1</td>
</tr>
<tr>
<td>73</td>
<td>7</td>
<td>16</td>
<td>1.51</td>
<td>2.13E-13</td>
<td>0.055</td>
<td>IGHG1</td>
</tr>
<tr>
<td>42</td>
<td>5</td>
<td>16</td>
<td>2.07</td>
<td>1.29E-12</td>
<td>0</td>
<td>IGHG1</td>
</tr>
<tr>
<td>73</td>
<td>5</td>
<td>15*</td>
<td>1.41</td>
<td>7.42E-11</td>
<td>0.082</td>
<td>IGHG1</td>
</tr>
<tr>
<td>33</td>
<td>5</td>
<td>14*</td>
<td>1.52</td>
<td>7.69E-11</td>
<td>0</td>
<td>IGHG1</td>
</tr>
<tr>
<td>17</td>
<td>5</td>
<td>13</td>
<td>1.18</td>
<td>1.14E-11</td>
<td>0.059</td>
<td>IGHG1</td>
</tr>
<tr>
<td>16</td>
<td>6</td>
<td>13</td>
<td>17.63</td>
<td>1.01E-11</td>
<td>0</td>
<td>IGHM</td>
</tr>
<tr>
<td>22</td>
<td>5</td>
<td>11</td>
<td>25.59</td>
<td>5.20E-10</td>
<td>0.091</td>
<td>IGHA1</td>
</tr>
</tbody>
</table>
HOMOLOGY TO APPROVED KDR MAB TRANSLATES TO TARGET BINDING

Ramucirumab is a VEGFR2 (KDR) antibody, approved for treatment of solid tumors

One of Alchemab’s convergent clusters is highly homologous to ramucirumab

- Assay shows binding for multiple antibodies from clonotype 41671 to KDR

![Graph showing binding data for different antibodies and controls](image-url)
Neuro degeneration
CASE STUDY
DEFINING RESILIENCE IN AD

MEASURE OF RESILIENCE

- Mini Mental State exam
- Cognitive function
- ApoE status and other genetic mutations
- B-amyloid (CSF/plaques)
- Tau (CSF/plaques)
- Additional biomarkers (NfL)

RESILIENT SAMPLE COHORTS

- Commercial study
  - 100+ cohort EPAD*

EPAD

100+ cohort

Report of Extraordinary individual with Presenilin (PSEN1) mutation but considerable delay of cognitive decline, with high b-amyloid plaque load but low Tau


* European Prevention of Alzheimer’s Dementia
ALZHEIMER’S DISEASE PILOT STUDY

- Longitudinal study over 2 years (visits every 6 months)
- Mini Mental State Exam score used for assessment of disease progression
- Starting range 15-24 out of a maximum score of 30

### Study enrolment
- 681

### Diagnosis of AD
- 333

### Age at entry >59
- 268

### Minimum 4 visits
- 63

### Δ MMSE score visits 1/4

<table>
<thead>
<tr>
<th>Progressor Type</th>
<th>Number</th>
</tr>
</thead>
<tbody>
<tr>
<td>Slow progressor</td>
<td>13</td>
</tr>
<tr>
<td>Average progressor</td>
<td>12</td>
</tr>
<tr>
<td>Fast progressor</td>
<td>12</td>
</tr>
<tr>
<td>Matched control</td>
<td>12</td>
</tr>
</tbody>
</table>

Elite controllers
IDENTIFICATION OF AD CONVERGENT CLONES

Sequence data from 30 Alzheimer’s patients was triaged to select 38 antibodies according to the following workflow:

- **CLONOTYPING**: sequences with near identical V, J, and CDR3 length are clustered
- **EXCLUSIVITY**: Clonotypes chosen which are exclusive to resilient group and not seen in healthy controls
- **EXPRESSION FILTER**: Sequence seen more than once per sample
- **MUTATION FILTER**: Filter for clonotypes undergoing active selection via somatic hypermutation

Exclusively convergent clonotypes have a highly skewed VH germline gene distribution which indicates disease specific activation
ALZHEIMER’S DISEASE ANTIBODY SELECTION

- **SEQUENCES**: 6,300,000
- **CONVERGENT CLUSTERS**: ~200
- **ANTIBODIES TRIAGED FOR PRODUCTION**: 38
- **ANTIBODY HITS**:
  - 8 antibodies bind immune cells
  - 2 antibodies bind neuronal cells
  - 3 antibodies bind astrocytes

Further analysis ongoing

**Additional Resources**
- Immunophenotyping
- Neuronal cell lines
- Human iPSC
AD CONVERGENT ANTIBODY SHOWING NEURONAL BINDING (TARGET UNKNOWN)

Alchemab Ab ATL 4527
VGLUT1 (pre-synaptic)
ALCHEMAB DATA CUBE: EXCEPTIONAL INSIGHT INTO BCR REPERTOIRES AND THEIR ASSOCIATION WITH DISEASE

- FULL STACK AND ADVANCED ANALYTICS PIPELINE

- DEEP LEARNING TOOLKIT:
  - Quality assessments
  - Clustering
  - Cross disease analysis
  - In silico light chain pairing
  - Patient stratification

...enabling a powerful next generation drug discovery engine
ANTIBERTA: ALCHEMABS’S ANTIBODY-SPECIFIC LANGUAGE MODEL

V Gene Family
- IGHV1
- IGHV2
- IGHV3
- IGHV4
- IGHV5
- IGHV6
- IGHV7

Number of Mutations
- 0: 24
- 8: 32
- 16: 40

Naive vs. Memory B Cell
- Naive
- Memory
CONVERGENCE ANALYSIS COULD BE PART OF THE NEXT WAVE OF ANTIBODY GENERATION

(A) Mouse hybridoma
- Immunization With targets
- Harvest Splenocytes Generate hybridomas
- Screening
- Chimerization
- CDR graft
- Mouse mAb

(B) Phage display
- scFv or Fab
- Phage-displayed Ab libraries
- Biopanning With targets (3-5 cycles)
- Screening
- Chimeric mAb
- Humanized mAb

(C) Transgenic mouse
- Immunization With targets
- Harvest Splenocytes Generate hybridomas
- Screening
- Constructing of Human IgG
- Human mAb

(D) Single B cell
- Immunization With targets
- Harvest Splenocytes Generate hybridomas
- Sort B cells with Labeled antigens
- Screening
- PCR, construct $V_H$ and $V_L$
- $V_{\beta}J_{\beta}C_{\beta}$
- $V_{\mu}D_{\mu}C_{\mu}$
- Human mAb

(E) B cell convergence
- Computation
- Convergence & triage
- Screening

Source: Journal of Biomedical Science, 21:1 (2020) - Adapted
OUR VISION: TO USE THE POWER OF THE HUMAN IMMUNE SYSTEM TO DISCOVER NEW MEDICINES AND DIAGNOSTICS

Discovering and developing protective, patient-originated therapeutic antibodies

Focus on protective antibody responses
Convergent in resilient individuals
Target agnostic approach
Advanced computational approaches