Global Investigation of Immune Repertoire Suggests Kawasaki Disease Has Infectious Cause

Ho-Chang Kuo, MD, PhD; Cheng-Tsun Pan, PhD; Ying-Hsien Huang, MD, PhD; Fu-Chen Huang, MD; Yeong-Shin Lin, PhD; Sung-Chou Li, PhD; Lien-Hung Huang, PhD

Background: Kawasaki disease (KD) severely threatens young children's health worldwide. The pathogenic mechanism of KD has not yet been solved, so there is still debate over whether KD is an infectious disease or an autoimmune disease.

Methods and Results: To solve this problem, an immune repertoire analysis of KD was conducted. We collected blood cell RNA samples and prepared them into amplicons with iRepertoire kits. The amplicons were sequenced and analyzed with the iRepertoire pipeline. We first identified KD-specific VJ and VDJ forms that had the potential to serve as biomarkers of KD. In addition, the KD-specific VDJ forms were contributed mostly by immunoglobulin G. The D50 value analysis showed that B-cell diversity in KD is decreased, suggesting unique immunoglobulins are produced in KD. Moreover, V, D and J segment usage in IgA, IgG and IgM was consistent with previous KD studies. Further comparison showed no difference in CDR3 peptide length between KD and fever controls (subjects with fever but not diagnosed as KD), indicating KD had B-cell selection phenomenon that has a non-autoimmune pattern. The comparison of amino acid usage of the CDR3 region demonstrated a preference for hydrophilic amino acids in KD.

Conclusions: The results of D50 value, VDJ usage and CDR3 peptide length analyses suggested the characteristics of infectious disease for KD.

Key Words: B cell; Immune repertoire; Immunoglobulin heavy chain; Kawasaki disease; VDJ form

Muco-cutaneous lymph node syndrome is an acute systemic vasculitis that predominantly affects the children under 5 years of age. It is also called Kawasaki disease (KD) because it was first described by Tomisaku Kawasaki in 1967.1 During the acute phase, coronary artery lesions (CALs) occur in less than 5% of all KD patients.2 and although KD has been studied for almost 50 years, the pathogenic mechanism of KD remains unclear. In KD patients, inflammatory cells may secrete pro-inflammatory cytokines and chemokines to activate the immune system.3–5 triggering endothelial cells to express adhesion molecules and to interact with inflammatory cells. As a result, inflammatory cells, including T lymphocytes, B lymphocytes, macrophages, neutrophils and plasma cells, infiltrate the coronary artery, leading to CALs or coronary artery aneurysms.6–9

In the mammalian immune system, B cells secrete immunoglobulins (Ig) and T cells produce T-cell receptors (TCR). The immune repertoire is defined as the number of forms of Ig or TCR one organism can produce. In both Ig and TCR, there are protein domains responsible for antigen recognition and binding. The DNA fragment encoding these protein domains can be classified into V, D, J and C regions (no D region for TCR). Taking the heavy chain of human B-cell Ig for example, there are approximately 80 hIGHV (human immunoglobulin heavy chain V), 30 hIGHD (D) and 6 hIGHJ (J) regions. During the maturation of lymphocytes, the B cells and T cells undergo VDJ recombination, through which the DNA fragments are trimmed out and only 1 copy of the V, D and J regions are ligated into the mature VDJ forms (VDJ gene mRNAs) in mature lymphocytes. As a result, B cells and T cells produce millions of types of Ig and TCR, respectively, greatly enhancing immune diversity.

Although KD has been studied for almost half a century, there is not yet a consensus on whether KD is an infectious disease or an autoimmune disease. The former is caused by a specific pathogen, while the latter is not. If KD is an infection, it should have a specific pathogen and during the acute phase of the disease the immune system of the affected...
patient should produce specific Ig or TCR that recognize and resist the KD pathogen. Therefore, by comparing the immune repertoire patterns of control subjects and KD patients, KD-specific VDJ form should be detected.

Globally investigating the immune repertoire of an individual was not possible until the invention of next-generation sequencing (NGS) technologies. With PCR primer sets specific for different VDJ forms, researchers can amplify desired forms and prepare them into amplicons, followed by NGS. By further mapping the NGS reads back to VDJ libraries, researchers can determine the expression abundance of detected VDJ forms. In this study, we hypothesized that we could determine whether KD is an infection or an autoimmune disease by comparing the immune repertoire of KD and control subjects. Further, if KD is an infectious disease, we hypothesized that the KD-specific VDJ forms can serve as KD biomarkers.

**Methods**

**Ethics, Consent, Permissions and Subject Enrollment**

We enrolled 40 subjects based on Institutional Review Board approval (IRB no. 201601004B0C101) by Chang Gung Memorial Hospital, Kaohsiung, Taiwan. Among the enrolled subjects, 20 were fever controls (FC), denoting subjects with fever but not diagnosed as KD, and the remaining 20 subjects were KD patients in the acute phase before intravenous Ig G (IVIG) treatment. To simplify the overall experiment design and results, we did not include incomplete KD subjects. In addition, the KD subjects with IVIG resistance or CALs were excluded. All subjects signed the informed consent form in accordance with the IRB.

**White Blood Cells (WBC) Enrichment and RNA Collection**

Total WBCs were enriched from whole blood by lysing red blood cells and further suspending them in phosphate-buffered saline. Following WBC enrichment, total RNA was isolated by mirVana™ miRNA Isolation Kit (Ambion, CA, USA), according to the manufacturer’s protocol. In addition, all RNA samples were analyzed with Bioanalyzer to guarantee that their RNA integrity number (RIN) values were not less than 8.

**VDJ Form Capture**

The collected RNA samples were amplified by HBHI-M reagent (Patent 7999092, iRepertoire, Inc. www.irepertoire.com), which is designed to capture the heavy chain of Ig produced by B cells. Each sequencing run contained 10 pooled libraries, comprising 5 FC and 5 KD RNA samples distinguished with multiplex barcodes. The sequencing job was done on Illumina MiSeq with the following setting: 250-cycle pair-end read. All preparation and experimental procedures followed the official protocols.

**Sequencing Data Analysis**

Raw reads were uploaded to and analyzed by the iRepertoire pipeline (iRepertoire, Inc.) with the following parameters specified: paired-end stitching, default SMART filters and V(D)JC gene mapping with IMGT® database. The output data downloaded from iRweb consisted of the mapped V(D)JC genes and the translated CDR3 peptides with counts, which were then normalized to 1 million for each library in order to neutralize unequal sequencing yields among libraries. D50 was calculated as the percent of dominant and unique V(D)J form that accounted for the cumulative 50% of total counts. Each V(D)J form was further classified into Ig types, IgA, IgD, IgE, IgM and IgG, by their corresponding mapped C gene.

**Results**

**Subjects’ Characteristics**

To answer our hypotheses, we interrogated the immune repertoire of KD in both patients and control subjects. Table shows that the majority of the FC subjects had either an upper or lower respiratory tract infection. All the KD patients had fever for more than 5 days and had at least 4 of 5 symptoms listed by the American Heart Association as diagnostic criteria.

**Implications of D50 Value**

B cells produce specific Ig under the stimulus of specific antigens. Hence, B-cell diversity depends on the types of Ig produced in an individual. The most important and polymorphic region where Ig interacts with antigen is the CDR3 region, organized by VDJ gene recombination. By investigating CDR3 variability, the immune repertoire of an individual when reacting to antigens can be interrogated. To estimate the global change in B-cell diversity between the FC and KD subjects, the D50 value was used to represent...
The diversity of antibodies in each sample. The D50 is the percentage of unique VDJ forms contributing to the top 50% of cumulative expressed NGS reads (also called counts). The more diverse the VDJ forms are within a sample, the closer to 50 is the value, in which the top 50% reads are equally contributed by 50% of unique VDJ forms.

The D50 value was calculated by in-house scripts adapted from the iRweb equation.11 As shown in Figure 1, D50 was globally smaller in the KD group than in the FC group (All), demonstrating that the VDJ pattern in KD was

**Figure 1.** Comparison of D50 value between fever control (FC) and Kawasaki disease (KD) samples. Data are presented as mean±SEM. Although D50 values between the 2 sets were not statistically significant, the tendency can be observed. With more samples included, statistical significance can be expected.

**Figure 2.** Two-dimension abundance of detected VJ forms between fever control (FC) and Kawasaki disease (KD) samples. By default, the iRepertoire pipeline reports the VJ patterns rather than the VDJ ones. VJ forms for FC457 (control subject, Top) and KD783 (KD subject, Middle), (Lower) Four significant VJ forms are shown. **P<0.001. Y axis denotes the unit of count per million.
samples. To date, the clear mechanism of class switching in KD remains unknown and deserves more attention.

**Identification of KD-Specific VJ and VDJ Forms**

After analysis by iRweb pipeline, the expression abundance of detected VJ forms was determined. Without considering the diversity of the hIGHD region, there were 288 VJ forms detected among the 40 samples and the expression abundance was determined. In representative FC and KD samples it was obvious that the combination V4-59+J4 form had almost 2-fold higher expression abundance in the FC than in the KD sample (Figure 2). To identify all differentially expressed VJ forms, we conducted a t-test and identified 20 significant VJ forms (P<0.05). Of these, 4 dominated by few VDJ forms and was less diverse. Therefore, most of the KD antibodies were contributed to by fewer unique VDJ forms, implying a reaction with a certain type of antigen in KD. Figure 1 also shows that some VDJ forms in KD had higher expression, diluting the D50 value, especially in IgM.

Ig class switching from IgM and IgD to IgG, IgE and IgA is a major characteristic of B-cell immunology. Both endogenous cytokine-dependent and antigen-induced switching have been reported. In this study, we found 34 VDJ forms that showed Ig type switching at 100% from IgM to IgA in KD groups (Supplementary Table 1). Among the 34 VDJ forms, 12 were expressed in FC samples but without class switching and 16 were not expressed in FC samples. To date, the clear mechanism of class switching in KD remains unknown and deserves more attention.

**Figure 3.** Significant VDJ forms between fever control (FC) and Kawasaki disease (KD) samples. All VDJ forms shown are P<0.01. Among the illustrated examples, 7 VDJ forms are more abundant in FC samples and 4 show a higher level in KD. Y axis denotes the unit of count per million.
are illustrated (Figure 2) and they were all more abundant in FC than in KD. In addition to the t-test, these significant VJ forms were also evaluated by a non-parametric test, the Wilcoxon rank-sum test. All of the significant VJ forms by t-test were also significant by Wilcoxon rank-sum test.

We further divided the NGS reads mapped back to the hiGHD regions into different hiGHD forms, identifying all VDJ forms. As a result, 6,429 VDJ forms were detected (Supplementary Table 2) among all 40 samples and 141 of them were significant by t-test (P<0.05). Figure 3 illustrates the 11 VDJ forms with P<0.01. Among these, 7 VDJ forms were more abundant in FC samples and the remaining 4 had higher levels in KD. In our previous study, we identified miRNA-based KD biomarkers and developed a KD diagnosis model.17 From our present study, the identified KD-specific VDJ forms may also serve as KD biomarkers, enabling development of a KD diagnosis model with increased sample size.

In humans, Ig can be classified into IgA, IgD, IgE, IgG and IgM types. In addition to statistical significance, we were also interested in which Ig types accounted for the statistical significance of the VDJ forms. In fact, the PCR primer sets to capture VDJ form also covered the C region (following the VDJ region). The C region has different IgA, IgD, IgE, IgG and IgM types, so the NGS reads covering the C region can be used to classify the Ig types. As a result, in addition to the overall VDJ forms, different Ig-based VDJ forms can also be determined. We determined the expression abundance of Ig-based VDJ forms. Figure 4 illustrates 9 significant Ig-based VDJ forms, 7 of which belonged to IgG. The current treatment of KD is administration of IVIG. Whether these 2 factors are biologically correlated or not deserves further investigation.

**Global and Ig-Specific V, D and J Usage**

So far, we observed KD-specific VJ and VDJ forms with splitting of V, D and J into subtypes. We further investigated V, D and J usage in the KD and FC samples, considering the major types of V, D and J. VDJ usage may reflect particular diseases.18–21 Figure 5 shows that, globally
Figure 5. VDJ segment usage of the overall and 5 immunoglobulin (Ig) types. The data are presented as mean±SD; P<0.05 indicates statistical significance, as determined by unpaired Student’s t-test. FC, fever control; KD, Kawasaki disease.
Immune Repertoire of KD

Immune Repertoire of KD

Therefore, we determined the amino acid usage in the CDR3 region, globally and Ig-specifically. Figure 6 shows that the hydrophilic amino acid preference was different between the FC and KD samples, especially in IgA and IgG. In particular, the tyrosine and alanine usage globally increased in KD. Further analysis found tyrosine usage increased in IgA, IgG and IgE. In addition, alanine usage increased in IgA in KD compared with FC samples. We also found tryptophan usage increased in IgG and IgE in KD. These findings implied B-cell selection to produce unique Ig in KD, a characteristic of infectious disease.

Comparison of Peptide Length in the CDR3 Region
The CDR3 length of IGH is associated with self-reactive or autoimmunity. The IGH CDR3 length is longer in early B-cell precursors than in mature peripheral B cells. The length of IGH CDR3 is also relatively long in polyreactive monoclonal antibodies. To identify the effect of CDR3 selection, we analyzed the CDR3 peptide length in overall Ig and in individual Ig types. Supplementary Figure shows there were no differences in CDR3 peptide length in both the overall and 5 individual Ig types. This result indicated that the self-reactive antibodies and polyreactive antibodies were selected and eliminated during B-cell maturation in KD, implying a non-autoimmunity pattern for both the FC and KD samples. In fact, all the FC samples were infectious disease, which was consistent with the analysis result for CDR3 peptide length.

Comparison of Amino Acid Usage in the CDR3 Region
The CDR3 region is critical for specific antigen recognition and the amino acid sequence in the CDR3 region may influence the binding between antibody and antigen. Positively charged amino acid enrichment is associated with anti-DNA antibodies in autoimmune disease. Therefore, we determined the amino acid usage in the CDR3 region, globally and Ig-specifically. Figure 6 shows that the hydrophilic amino acid preference was different between the FC and KD samples, especially in IgA and IgG. In particular, the tyrosine and alanine usage globally increased in KD. Further analysis found tyrosine usage increased in IgA, IgG and IgE. In addition, alanine usage increased in IgA in KD compared with FC samples. We also found tryptophan usage increased in IgG and IgE in KD. These findings implied B-cell selection to produce unique Ig in KD, a characteristic of infectious disease.

Discussion
In this study, we analyzed the global and Ig-specific V, D and J usage in KD and FC subjects. In the acute phase of KD, IgA-secreting plasma cells infiltrate the vascular wall into the media. The serum level of IgA antibodies also increases significantly, indicating that the IgA immune system is activated in KD. Our data showed that particular VDJ segment usage increased in IgA. This evidence implies an antigen-specific immune response in KD. Furthermore, IgG and IgM plasma cells are present in the coronary artery in the acute phase of KD, and the serum levels of IgG and IgM also markedly increase in KD.

Epidemiological and clinical evidence points to KD having an infectious etiology. KD is characterized by a seasonal peak in many areas and cities. In the acute phase of KD, the total WBC count is increased: CD14+ monocytes and CD19+ B cells are significantly increased, but not T lymphocytes or NK cells. The CD3+, CD4+, CD8+ T lymphocytes and CD56+ NK cells are decreased in peripheral blood. Immunochemical studies have shown that neutrophils infiltrate in CALs in the early stage of KD. Monocytes/macrophages are the majority of inflammatory
cells in CALs. CD3+ T lymphocytes and CD19+ B lymphocytes also infiltrate the CALs of KD.6 This evidence indicates an important role for the immune system in KD.

In this study, we used NGS technology to analyze the expressed B-cell receptor genes in the immune repertoire of KD. We observed specific VDJ forms differentially expressed in the FC and KD samples. The future analysis showed a discrepancy between the FC and KD samples in the expression abundance of some Ig-based VDJ forms: 7

showed a discrepancy between the FC and KD samples in expressed in the FC and KD samples. The future analysis indicates an important role for the immune system in KD.

Previous studies also indicated that a mucosal immune response was involved in KD. The serum levels of anti-lipid IgA are significantly elevated in KD.7 Increased serum levels of IgA, IgM and C3 in KD patients have been observed,8 and the numbers of IgA, IgM and IgG-secreting cells increases in acute-phase KD patients.9

Immunohistochemical studies have shown that IgA, IgM and IgG plasma cells infiltrate the vascular wall in KD.9 Not only vascular tissue, IgA plasma cells also infiltrate the coronary artery, pancreas and kidney. Rowley et al sequenced the VDJ junctions of a genes and concluded that the vascular IgA response in acute KD was oligoclonal.10 Consistent with that, our analysis showed that particular VDJ segment usage was highly diverse in IgA.

Conclusions
It is still a debate whether KD is an infectious or autoimmune disease. In this study, we investigated the immune repertoire of the IGH in KD and FC samples. We first identified KD-specific VJ and VDJ forms that have the potential to serve as biomarkers for KD diagnosis. In addition, the analysis results of D50 value, VDJ usage and CDR3 peptide length suggested the characteristics of infectious disease for KD, implying KD is an infection not an autoimmune disease.

Declarations
Competing Interests: The authors declared there was no conflict of interests.

Consent for Publication: The subjects contributing clinical samples to this study (or their guardians) fully understood the purposes of this study and signed the informed consent voluntarily.

Funding: Publication of this article has been funded by the grants from Chang Gung Memorial Hospital (CMRPG5F1941 and CORPGF90011-3).

Authors’ Contributions: LHH conducted the experiments. CTP, YSL and SCL were responsible for data analysis. FCH and YHH were responsible for clinical sample collection and experiment design. HCK and SCL were responsible for data analysis. FCH and YHH were responsible for technical support. We thank the Genomics & Proteomics Core Laboratory, Department of Medical Research, Kaohsiung Chang Gung Memorial Hospital for technical support.

Acknowledgment
We thank the Genomics & Proteomics Core Laboratory, Department of Medical Research, Kaohsiung Chang Gung Memorial Hospital for technical support.

References
27. Takeshita S, Kawanishi H, Shimazu T, Yoshida M, Seike I. Increased production of serum IgA-class antibody to lipid A in...


**Supplementary Files**