博士論文（要約）

Analysis of brownbanded bamboo shark (*Chiloscyllium punctatum*) immunoglobulin novel antigen receptor (IgNAR) as a potential immunotherapeutic

抗体医薬として期待されるイヌザメ（*Chiloscyllium punctatum*）免疫グロブリン新規抗原受容体 IgNAR の解析

Dehinga Prasadi Nayanashani De Silva

デヒンガ　プラサディ　ナヤナシャニ　デ　シルワ

Analysis of brownbanded bamboo shark (*Chiloscyllium punctatum*) immunoglobulin novel antigen receptor (IgNAR) as a potential immunotherapeutic

抗体医薬として期待されるイヌザメ(*Chiloscyllium punctatum*)免疫グロブリン新規抗原受容体 IgNAR の解析

A thesis submitted to

Graduate School of Agricultural and Life Sciences, The University of Tokyo

in partial fulfillment of the degree of

Doctor of Philosophy

in Department of Aquatic Bioscience

By

Dehinga Prasadi Nayanashani De Silva

March, 2020

DECLARATION

I, Dehinga Prasadi Nayanashani De Silva, hereby declare that the thesis entitled **"Analysis of brownbanded bamboo shark (*Chiloscyllium punctatum*) immunoglobulin novel antigen receptor (IgNAR) as a potential immunotherapeutic"** is an authentic record of the work done by me as a Doctoral student of Department of Aquatic Bioscience, Graduate School of Agricultural and Life Sciences, the University of Tokyo and that no part thereof has been presented for the award of any degree, diploma or any other similar title offered by other education institutions.

D.P.N. De Silva

Laboratory of Aquatic Molecular biology and Biotechnology,

Department of Aquatic Bioscience,

Graduate School of Agricultural and Life Sciences,

The University of Tokyo

March, 2020

DEDICATION

To

My mother,

For giving me love, care & strength,

To

My daughters,

For making me a mother, to give love & endure hardships

Acknowledgements

I would like to express my sincere gratitude to Prof. Shuichi Asakawa, who kindly accepted to supervise me and guide me towards Ph.D. Your advices and valuable critics always cleared the path for during the research. With deep appreciation I would like to acknowledge the efforts made to make this research a fruitful one by Prof. Shigeharu Kinoshita and Prof. Susumu Mitsuyama. I am indebted to Prof. Yoshitake for his generous support on bioinformatic analysis. As a beginner I was trained and guided by Dr. Tan and Dr. Igarashi, whom I would like to convey my sincere gratitude. It is my great pleasure to extend my heartfelt thanks to Prof. Sho Hosoya, Mr. Mizuno and Mr. Fujitha for their enormous support during sample collection. Without them and the staff of Aquatic Research Center, Hamamathsu, sampling would have been a nightmare for me. I am very grateful to the staff members Mr. Onda Kiyoko and Mr. Inoda Takahiro at Kaiyukan, Osaka for providing bamboo sharks for the experiment.

I must not forget to be grateful to Prof. Takeshi Omasa for his kind acceptance to visit his laboratory at Osaka University to learn CHO cell culturing procedure and also providing mammalian expression vector made in his lab. I am very thankful to Dr. Jana, who taught me how to perform CHO culture and my heartiest gratitude goes to the staff and students at the Division of Science and Biotechnology, Graduate School of Engineering, Osaka University.

I would like to express my deepest gratitude to Prof. Hidehiro Kondo, for allowing me to use their Illumina sequencer and his kind guidance on this research. I am grateful to Prof. Ikuo Hirono and the staff members of the Laboratory of Genome Science,

Tokyo University of Marine Science and Technology for their kind support to conduct this research smoothly. My sincere gratitude goes to Prof. S. Watabe and Prof. S. Reza of Kitasato University as well for allowing me to use their facility.

I am very grateful to Prof. Hideki Ushio, for his valuable thoughts and sharing resources throughout the period of my research and Dr. Misako Nakaya, who has been helping me in numerous ways. My sincere gratitude goes to all the professors and staff of the Department of Aquatic Bioscience and Graduate School of Agricultural and Life Sciences for helping me in enormous ways during my study period at The University of Tokyo. My research life became full of unforgettable memories thanks to the colleagues at the Laboratory of Aquatic Molecular Biology and Biotechnology.

My study in Japan was funded by the Ministry of Education, Culture, Sports, Science and Technology, Monbukagakusho scholarship and I am very grateful for their support. Also I am very grateful to Uva Wellassa University and the University Grants Commission of Sri Lanka for providing study leave to pursue doctoral degree.

I am very fortunate to have been raised by a strong woman, who is my mother. Despite of her illnesses, she came to Japan to help me through. I am thankful to my dear father and only sister for their love and encouragement. I am forever grateful to my husband, who did lot of sacrifices to make me reach success. Finally, I am so pleased to say my heartiest thanks to my two little angels for their arrival to my world during the period of this research.

Table of contents

	Page
Declaration	I
Dedication	II
Acknowledgement	III
Table of contents	V
Abbreviations	IX
List of tables	XI
List of figures	XIII
Abstract	XVIII
List of publications	XXV

Chapter 1: General introduction

1.1 General background	2
1.2 Objectives of the study	12
1.3 Outline of the thesis	13

Chapter 2: Characterization of brownbanded bamboo shark (*Chiloscyllium punctatum*) IgNAR constant domains

Abstract 16

2.1 Introduction 17

2.2 Methods 21

2.3 Results 26

2.4 Discussion 36

Chapter 3: Diversity of brownbanded bamboo shark (*Chiloscyllium punctatum*) vNAR in response to antigen stimulation

Abstract 42

3.1 Introduction 43

3.2 Methods 46

3.3 Results 53

3.4 Discussion 67

Chapter 4: Transcriptome analysis provide insights to gene expression profile of lymphoid cells in brownbanded bamboo shark (*Chiloscyllium punctatum*)

Abstract 72

4.1 Introduction 73

4.2 Methods 76

4.3 Results 78

4.4 Discussion 88

Chapter 5: Selection of the antigen specific vNAR candidate gene by expression analysis

Abstract 91

5.1 Introduction 92

5.2 Methods 95

5.3 Results 108

5.4 Discussion 120

Chapter 6: General discussion

6.1 General discussion 124

6.2 Future prospective 128

References **132**

Annexure 01 **145**

Appendix 1 145

Appendix 2 162

Appendix 3 166

Appendix 4 173

Annexure 02 **177**

Abbreviations

Ab – Antibody

ADC - antibody-dependent cytotoxicity

APCs – Antigen presenting cells

BSA – Bovine serum albumin

C – Constant region

CDC- complement-dependent cytotoxicity

CDR – Complementarity determining region

CH – Constant heavy chain CHO – Chinese hamster ovary

DC - dendritic cells

DEGs – Differentially expressed genes

DNA - Deoxyribonucleic acid

DPBS - Dulbeccos' phosphate buffered saline

E-PBS - elasmobranch-phosphate buffered saline

FBS – Foetal bovine serum

FR – Framework region

H – Heavy chain

HEL – Hen egg lysozyme

HSA – Human serum albumin

IgNAR – Immunoglobulin novel antigen receptor

Igs - Immunoglobulins

L – Light chain

LO – Leydig organ

mAb – Monoclonal antibody

NGS – Next generation sequencing

PBL - peripheral blood leukocytes

PCR – Polymerase chain reaction

RAG - recombination activation gene

RNA-Seq – RNA sequencing

SOB – Super optimal broth

TdT - terminal deoxynucleotidyl transferase

V – Variable region

VHH – Nanobody from Camelids

vNAR - IgNAR variable domain

List of tables

Page

2-1 Cartilaginous fish IgNAR sequence NCBI accession numbers used in the local database to identify *C.punctatum* vNAR 25

2-2 Summary of the brownbanded bamboo shark transcriptome library indicating the sequence count and lengths before and after assembly by the Trinity or PEAR software 28

2-3 Minor variation frequencies and respective amino acid changes in IgNAR constant heavy chain C1 domain 29

2-4 Comparison of the bamboo shark IgNAR constant heavy chain protein sequence using the NCBI blastp with published cartilaginous species IgNAR sequences 31

3-1 Immunization of HEL: times of exposure of shark with their identification number (ID) 46

3-2 KAPA HiFi HotStart PCR protocol reagents and thermocycler protocol 48

3-3 Summary of the PEAR assembled cDNA database of *C. punctatum* vNAR 54

3-4 Number of variations in CDR1 and CDR3 amino acid sequences (unique domains) observed in four individual brownbanded bamboo sharks according to the level of antigen exposure 59

3-5 Presence of cysteine residue in different domains of variable region which was used for classification of IgNAR 63

4-1 Illumina sequencing reads of the control and immunized samples 78

5-1 Application of shark vNAR against various antigens 93

5-2 Amino acid residues from the candidate anti-HEL vNAR gene CDR1 and CDR3 domains which form contacts with HEL 113

List of figures

		Page
1-1	The T and B cell activation and function	3
1-2	Structure of IgG showing heavy and light chains and their specific sites	5
1-3	Evolution of immunoglobulins in vertebrates and development of their characteristic features	6
1-4	Illustration of immune tissues and organs in shark	7
1-5	Translocon and cluster organization of immunoglobulin genes among species	8
1-6	pEHX1.1 original vector map	11
2-1	Structure of IgNAR showing variable domains (CDR1, CDR3, HV2 and HV4) and five constant domains	18
2-2	Classification of IgNAR based on the non-canonical cystein residue in variable region	19
2-3	**a.** Internal anatomy of shark	21
	b. Collection of blood from the caudal vein of anaesthetized brownbanded bamboo shark	22
2-4	Quality of total RNA extracted from brownbanded bamboo shark PBL analyzed by Agilant tapestation 2200	26
2-5	FastQC per base quality scores of brownbanded bamboo shark raw sequence (forward strands) obtained from Illumina MiSeq.	27
2-6	Brownbanded bamboo shark IgNAR heavy chain C1 domain sequences indicating the presence of two IgNAR types	30
2-7	Phylogenetic analysis of the IgNAR heavy chain constant region complete	

sequence of the brownbanded bamboo shark (IgNAR-1 CH and IgNAR-2 CH) compared to other cartilaginous fish species 32

2-8 The brownbanded bamboo shark IgNAR (*C. punctatum* IgNAR-1 CH and C. punctatum IgNAR-2 CH) constant region amino acid alignment compared to other cartilaginous fish species 34

2-9 Amino acid sequence alignment of *C. punctatum* and other cartilaginous fish CH domains to Human IgG constant domains 35

3-1 Brief explanation on affinity maturation of antibody and its coding sequence 43

3-2 Three types of primers designed for the Illumina tailed PCR (PCR 0, PCR 1 and PCR 2) for amplification of vNAR gene 47

3-3 Schematic illustration of the ELISA designed for the detection of bamboo shark HEL specific IgNAR 51

3-4 Gel electrophoresis result of PCR-0, PCR-1 and PCR-2, amplifying brownbanded bamboo shark vNAR region for Illumina sequencing using KAPA HiFi HotStart 2x DNA amplification kit 53

3-5 Individual sample amplicon frequencies derived from Illumina Miseq sequencing of brownbanded bamboo shark vNAR region 55

3-6 Sequence length distribution of samples before (a) and after (b) immunization 57

3-7 Categorization of sequences of individual samples based on the functionality of immunoglobulin domains of vNAR region by IMGT/V-High Quest 58

3-8 CDR1 and CDR3 domains number of variable sequence diversity in individual brownbanded bamboo sharks at different antigen exposure levels 59

3-9 Expression level of cysteine among vNAR domains 61

3-10 Percentage of sequences according to the number of cysteines expressed in CDR1 and CDR3 domains of brownbanded bamboo shark vNAR 62

3-11 Most abundant patterns of cysteine availability in variable domains including three types IgNAR types 64

3-12 Brownbanded bamboo shark vNAR sequences with most diverse patterns in comparison with nurse shark vNAR 64

3-13 Plasma protein and IgNAR antibody analysis by SDS-Page and Western blotting 65

3-14 Analysis of the relative binding of immunized plasma IgNAR in comparison to their unimmunized control against HEL antigen 66

4-1 The number of DEGs before and after exposure to HEL by three adult bamboo sharks 79

4-2 Comparison of the change of DEGs among three sharks 80

4-3 MA plot showing the up and down regulated genes of PBLs before and after exposure to antigen 81

4-4 Top 15 GO terms of A. down-regulated genes and B. Up-regulated genes of immunized shark 82

4-5 Enrichr analysis of genes expressed in PBLs after exposure to the antigen 83

4-6 GO terms of up regulated genes from antigen exposed PBL cells from three brownbanded bamboo sharks 84

4-7 Differentially expressed GO terms of spleen and Leydig organ from immunized shark 85

4-8 MA plot showing the comparison of DEGs of spleen and Leydig organ 86

4-9 Top 15 GO terms expressed in spleen and Leydig organ from immunized brownbanded bamboo shark 87

4-10 The comparison of genes expressed in spleen, LO and PBLs 87

5-1 Original vector and the insertion of nurse shark IgNAR gene sequence and IgG Fc region at restriction digestion site 96

5-2 A. Arrangement of restriction enzymes in Prof. Omasas' vector and two restriction enzymes used to remove nurse shark IgNAR 98

B. Shows that the brownbanded bamboo shark vNAR-C1 insert and deletion of nurse shark vNAR 98

5-3 Illustration of the different parts of the potential anti-HEL vNAR gene insert of brownbanded bamboo shark origin for the expression of nanobody 103

5-4 pEHX1.1-N.S. IgNAR-Fc vector map and random primer binding sites 105

5-5 Absorbance values from ELISA using plasma IgNAR from immunized shark-1 (left) and 5 (right) against HEL 107

5-6 Potential anti-HEL vNAR sequence abundance according to the number of HEL antigen exposure 110

5-7 Number of H bonds and contacts form between HEL and vNAR to stabilize Ag-Ab complex 111

5-8 Number of molecules interact between vNAR and HEL to form the Ag-Ab Complex 111

5-9 Ribbon (left) and hydrophobic (right) structures of anti-HEL vNAR candidate gene selected for expression 112

5-10 Amino acid residues that form contacts with antigen in CDR1 and CDR3

domains of brownbanded bamboo shark vNAR 113

5-11 Structural comparison for the Ag-Ab interactions (contacts) of the candidate anti-HEL vNAR from brownbanded bamboo shark vs nurse shark anti-HEL vNAR using UCSF Chimera 114

5-12 *E. coli* competent cells showing the growth of individual colonies containing pEHX1.1-N.S.IgNAR-Fc vector 115

5-13 Linearization of the plasmid vector pEHX1.1-N.S.IgNAR-Fc using two enzymes Hind III and Alo I 116

5-14 Agarose gel electrophoresis image of the bands obtained by using random primers on the pEHX1.1-N.S.IgNAR-Fc vector 116

5-15 Amplification of Brownbanded bamboo shark anti-HEL vNAR-C1 insert sequences 117

5-16 Constructed plasmid of bamboo shark vNAR insert and its single digestion product which shows ~3.5 kb and ~3 kb band sizes 118

5-17 PCR amplification of brownbanded bamboo shark vNAR construct vector sequence using random primers designed for modified vector 119

5-18 Growth of CHO-K1 cell line in EX-CELL media with 2% FBS 119

5-19 Different platforms currently using to isolate highly specific vNAR antibodies 120

6-1 Key results generated by the ongoing study on developing anti-BSA vNAR nanobody 129

6-2 Comparison of the binding ability of brownbanded bamboo shark anti-BSA vNAR with anti-BSA nanobody (5VNV) 130

Abstract

Immune system is important for animals to fight against invading pathogens and antigenic agents. It contains two parts mainly innate immunity and acquired immunity. During evolution, cartilaginous fish are the first jawed vertebrates having immunoglobulins (Igs) which is also called as antibodies. There are three types of Igs found in sharks namely, IgM, IgW and immunoglobulin novel antigen receptor (IgNAR). IgNAR was reported first by Greenburg et al., in 1995 as heavy chain homodimer. Since then many scientists focused on this antibody due to its advantageous characteristics in biomedical research.

Only two naturally occurring heavy chain only antibodies (lacking light chain) were reported, one from camelids called VHH and the other one is IgNAR from cartilaginous fish. Both of these molecules have smaller variable regions of which variable region of IgNAR (vNAR) considered as the smallest antibody like molecule in nature (~ 12 – 15 kDa). The reasons for the smaller size are due to lack of light chain and devoid of second complementary determining region (CDR). IgNAR was found to be stable in harsh environments such as high temperature and pH variations. This might be an adaptation of the Ig which circulates in shark blood with high ammonia concentration. The higher stability was believed to be due to the presence of non-canonical cysteine molecules present in the variable domains of vNAR.

Use of monoclonal antibodies on diagnostics and treatments is becoming more popular recently. Due to larger molecular size of the conventional antibodies (eg: IgG 150 kDa), it is difficult to reach the epitopes particularly in cellular spaces and cancer

cells. Hence scientists focused on engineering the antibodies to develop smaller, antigen specific and high affinity monoclonal antibodies. Variable region of shark IgNAR which comply with those features is being investigated to develop immunotherapeutics (anti-Ebola, anti-TNFα) and diagnostics (anti malaria AMA1, anti-Cholera). Various techniques have been used to develop these nanobodies such as immunization of shark and development of semi-synthetic antibodies.

There are four types of IgNARs classified to date based on the presence and location of cysteine residue in the variable region. Currently several sharks and rays have been investigated for their IgNAR gene sequence. Although they all have non-canonical cysteine residues, type I IgNAR was reported only from nurse shark (*Ginglymostoma cirratum*). Most of the investigators also performed immunizations and it revealed that different shark species have been responding differently against antigens. Hence, the aim of this study was to develop a novel platform to isolate antigen specific vNAR from brownbanded bamboo shark (*Chiloscyllium punctatum*) by *in vivo* affinity maturation.

1. Characterization of brownbanded bamboo shark IgNAR constant region

IgNAR has two regions namely, the variable region and constant region. Constant region consists of five domains C1 to C5. These domains help the integrity of the IgNAR antibody and also maintains flexibility at the hinge region during antigen binding. Study on brownbanded bamboo shark focused on the evolutionary compliance with the conventional antibodies particularly human Igs. In this study, brownbanded bamboo shark (*Chiloscyllium punctatum*) was selected as the model species to characterize the IgNAR constant region sequence by next generation sequencing (NGS) method. Peripheral blood leukocytes (PBL) were used to isolate mRNA and synthesized cDNA

for the library having 20 million raw reads (Illumina, MiSeq). Average sequence length after paired end assembly was 218.4 bp (35 - 301 bp) with total of 4.5 billion residues. Transcriptome analysis of complete constant region with five domains and conserved cysteine residues revealed the presence of two distinct types of IgNAR in brownbanded bamboo shark. The C1 domain alone displayed 13 unique sequences that might resemble the number of IgNAR CH gene clusters. Furthermore, phylogenetic analysis showed that the relationship with order Orectolobiformes, of which nurse shark (*Ginglymostoma cirratum)* and wobbegong shark (*Orectolobus maculatus*) as the highest similarity. The alignment of human IgG constant regions indicated the conserved nature of cysteine residues throughout the evolution of Igs. Through this study, understanding of the characteristics of IgNAR constant domains used to maintain its structural stability will become useful information for engineering monoclonal antibodies in the future.

2. Diverse response by vNAR of brownbanded bamboo shark against antigen

Shark vNAR is widely used in therapeutic and diagnostic research due to its smaller size which enables tissue penetration, simple architecture and higher stability of the molecule that helps antibody engineering easier and cost effective. Developing highly specific and higher affinity antibody is important to minimize adverse effects. Presence of cysteine in variable region is useful to maintain structural integrity and stabilize the antigen binding regions. Number of cysteine residues also play an important role in the formation of diverse antigen-binding surfaces. Therefore, studying the responses against antigen and screening cysteine expression in CDR domains is important prior to synthesis of the therapeutic or diagnostic nanobodies. This study aimed at understanding the diverse nature of vNAR from brownbanded bamboo shark and the individual response

against antigen at various time of exposures. Adult sharks were immunized to determine antigenic responses of vNAR by injecting hen egg lysozyme (HEL) and collected the peripheral blood leukocytes (PBL) for cDNA library synthesis. Three-step tailed PCR was designed and vNAR library was obtained using Illumina amplicon sequencing. The vNAR library with 72% productive sequences was consisted of 8.41 GB data (19.3 million sequences). More mutations were found in CDR-1 than CDR-3 which is typical for type II IgNAR when affinity maturation occurs due to somatic hypermutation. However, CDR1 and CDR3 variations were reduced as number of antigen exposures increased. Type II IgNAR sequences were abundant in only one shark and found different patterns expressed in different individuals. Interestingly, type I sequence found in one shark after first exposure to antigen at a very low expression level (1/100,000). In addition to nurse shark this is the first report of type I IgNAR from another species. There was an increase of IgNAR titer along with vNAR sequences obtained after HEL exposure. This study reveals the diversity of vNAR among individual sharks and also variability during the number of exposures. Hence it is suggestive of determining the highly antigen specific vNAR gene sequence prior to development of monoclonal antibodies from shark origin. This study shows the platform of synthesizing large cDNA library by amplicon sequencing and reduces the time, labour and cost of conventional cloning methods.

3. **Transcriptome analysis revealed overall positive immune response against antigen**

The transcriptome analysis and RNA-Seq data provides insights to gene expressions and reveals lot of information. Currently few shark species including nurse

shark tissues such as heart, spleen, kidney were studied. Genomic study on Elephant shark (*Callorhynchus milii*) showed that adaptive immune system of cartilaginous fish has T helper cells that were highly restricted and unconditional antigen-binding ability. Furthermore, it lacks CD4 and related cytokine receptors as well. However very few studies have been done on transcriptomic analysis of shark species and almost none of the studies investigated the immune response before and after antigen exposure in brownbanded bamboo shark. The aim of this study was to perform transcriptome analysis on the response of PBL from brownbanded bamboo shark against HEL. The RNA-seq library was prepared from PBL of pre-immunized (control) and immunized brownbanded bamboo sharks. Nearly 40 million clean sequence reads were obtained including 56 % immunized, 46 % control sequences. After antigen exposure, genes on negative regulation of Ig production was down regulated while regulation of complement activation, lectin pathway and regulation of cellular defense response were upregulated. Based on the transcriptomic analysis brownbanded bamboo shark shows great adaptive immune response against antigen which makes it a suitable candidate for immune research.

4. **Strategy to synthesize brownbanded bamboo shark vNAR based monoclonal antibody**

Development of monoclonal antibody based therapeutics is becoming popular these days having around 70 products in the market and over 100 in clinical trials. It is predicted to be increasing and advancing with the use of smaller, recombinant antibody fragments and engineered antibody variants such as nanobodies. Shark vNAR having interesting features fit for immunotherapy is a potential candidate and this study

performed to develop a platform to synthesize monoclonal antibody from brownbanded bamboo shark vNAR origin. Adult sharks were immunized with hen egg lysosome (HEL) and cDNA amplicon library prepared using mRNA from peripheral blood leukocytes. After paired end assembly unique sequences were listed based on the abundance and expression during prior and after immunization were compared. The sequences expressed after third immunization with higher abundance were selected and obtained anti-HEL vNAR model (nurse shark anti-HEL vNAR was used as template). Using Chimera (UCSF) software, strength of antigen binding by the number of contacts between antigen-antibody complex was determined. In order to synthesize anti-HEL nanobody, the highest affinity vNAR gene sequence was inserted into mammalian expression vector pEHX1.1-N.S.-IgNAR-Fc (from Omasa lab, Osaka University). The protocol has been developed to express the vNAR antibody using chicken hamster ovary (CHO) cell line. This novel platform for isolation and expression of antigen specific vNAR may be useful for the development of monoclonal nanobodies from shark origin at less cost and time compared to the conventional methods such as phage display.

5. Future research potential

Shark IgNAR was first reported in 1995 by Flajnik and co-workers. Since then many studies have been carried out on characterization of IgNAR in different shark species and synthesis of antigen specific vNAR by immunization or molecular engineering. However, there are a few drawbacks of using shark vNAR; firstly, due to the smaller size (~12 – 15 kDa) it is easily filtered through kidney glomeruli causing less retention time (or half-life), secondly, adverse effects due to rejection of the antibody from host. Hence this study focused on developing IgNAR against albumin, which is a

main plasma protein in mammals with higher molecular weight (66.5 kDa). A combination of HEL and bovine serum albumin (BSA) were injected to brownbanded bamboo sharks and synthesized a cDNA library with amplicons covering full-length vNAR gene sequences. Similar to the previous study with HEL, potential anti-BSA vNAR gene sequences were determined by using anti-HSA vNAR template from spiny dogfish. The target binding residues such as, Trp, Asp, Tyr and Pro were prominent in anti-BSA vNAR CDR3 sequence in brownbanded bamboo shark. According to ELISA, plasma IgNAR response against HEL became less compared to BSA. The sequences derived from this study will be useful for the expression of nanobodies using CHO cells in further research.

In conclusion, this research investigated brownbanded bamboo shark naiive IgNAR and immunized variations. The transcriptome analysis will provide insights to the antigenic response by PBLs. In addition, development of novel platform to synthesize full length vNAR amplicon library and screening highly specific gene sequences will be useful in developing nanobodies with minimal adverse effects in the future.

List of publications

Some contents of this thesis have been communicated as follows:

Journal article:

De Silva DPN, Tan E, Mizuno N, Hosoya S, Reza Md S, Watabe S, Kinoshita S and Asakawa S. (2018) Transcriptomic analysis of immunoglobulin novel antigen receptor (IgNAR) heavy chain constant domains of brownbanded bamboo shark (*Chiloscyllium punctatum*). *Fish and Shellfish Immunology*, 84:370-376

Conference proceedings and presentations:

De Silva, D.P.N., Hosoya, S., Mizumo, N., Kinoshita, S., and Asakawa, S. (2017) Transcriptome analyses on *in vivo* affinity maturation of immunoglobulin novel antigen receptor (IgNAR) from Brownbanded bamboo shark (*Chiloscyllium punctatum*) as an animal model for immunotherapeutics. Book of abstracts, 10th European zebrafish meeting, 3-7 July 2017. Budapest, Hungary. 458-459 .

De Silva, D.P.N., Hosoya, S., Mizumo, N., Kinoshita, S., Totsuka, M. and Asakawa, S. (2017) Determination of immunoglobulin novel antigen receptor (IgNAR) *in vivo* affinity maturation in brownbanded bamboo shark (*Chiloscyllium punctatum*). Proceedings of JSFS International Symposium, 22-24 September 2017. Shinagawa, Japan.

De Silva D.P.N., Mizuno N., Hosoya S., Kondo H., Yoshitake K., Kinoshita S. and Asakawa S. (2018) Nanobodies of shark origin: novel approach for cancer immunotherapy. Annual Research Conference, Sri Lanka Students Association in Japan. The Embassy of Sri Lanka, Tokyo, Japan.

Chapter 1

General introduction

1.1 General Background

1.1.1 Immune system

Immune system is a collection of proteins and cells which protects the body against pathogens (microbes), antigens, cancers and toxins. It can be divided into two main categories such as innate immune system and adaptive immune system. The innate immune system plays a role as the first line of defense. It is not specific and does not have memory. However, it acts immediately after exposure to an antigen. The adaptive immune system is specific for antigen and develops memory for the effective and rapid response during second exposure. There are two types of cells involved in adaptive immune system such as T cells and B cells. Although T cells require antigen presenting cells (APCs), B cells can recognize antigens directly by their antigen-binding receptors (figure 1.1). B cells produce antibodies (immunoglobulins) which binds to the surface receptors of the antigens (Warington et al., 2011).

The immunogenesis pathway is not a single cell process. It requires the combinatorial association of cells such as dendritic cells (DC), T lymphocytes and B lymphocytes. When an antigen entered to the body, DC recognizes the epitopes and endocytose it to present to the cell surface peptides of major histocompatibility complex (MHC) molecules. This stimulates T helper (T_h) cells to release cytokines which assist differentiation of the B lymphocytes in order to synthesize antibodies (McCullough & Summerfield 2005).

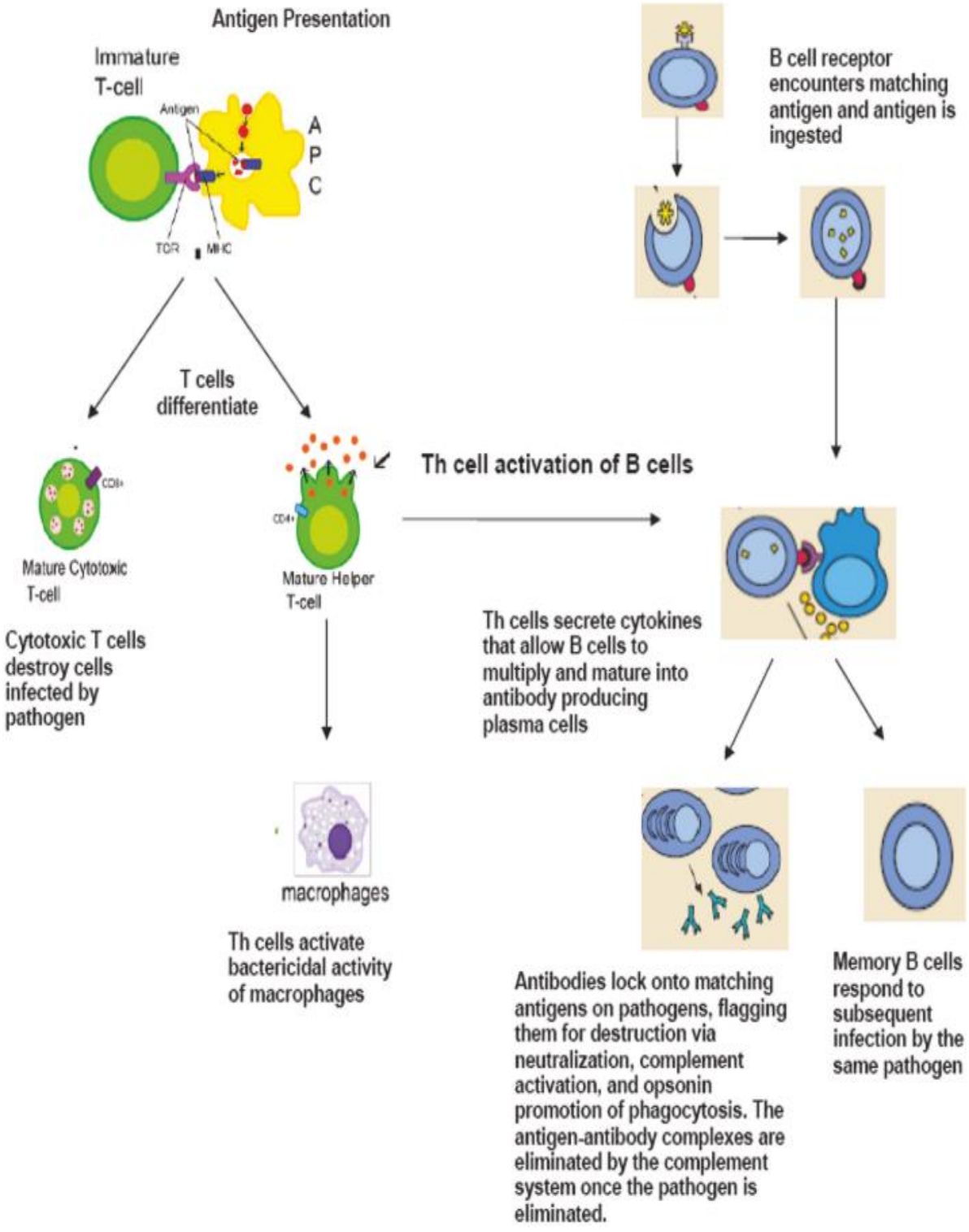


Figure 1.1. The T and B cell activation and function (source: Warrington et al., 2011)

1.1.2 Immunoglobulin structure and function

The history of the research on immunoglobulin goes way back to the year 1890, when von Behring and Kitasato first identified an agent which had the ability to neutralize the diphtheria toxin. Later on the so called terms of "Antibody" and "Antigen" were introduced by various scientists (Schroeder & Cavacini, 2010). By the advancement of electrophoresis separation methods, in 1939 two scientists named Tiselius and Kabat reported the globulin fraction; hence the terms "gamma-globulin", "immunoglobulin" was came in to picture. Along with the advancement of technology the investigations of structure and function of immunoglobulins were widely carried out in order to solve the complexity of their nature.

Antibody (Ab) or immunoglobulins (Igs) have the ability to bind to a number of different antigen types through adjustable binding due to alteration of deoxyribonucleic acid (DNA) arrangement in B lymphocytes. There are two types of Igs such as membrane bound form and soluble form. The membrane bound form acts as cell surface receptors, which convey the signal and activate cells after exposure to an antigen. The soluble forms are in plasma and able to bind with antigen individually and neutralize them at a distance (Nikolova et al. 2009). The basic structure of Igs includes two heavy chains (H) and two light chains (L) (figure 1.2). Both have variable (V) region with NH_2 - terminal and constant (C) region with COOH – terminal. (Williams & Barclay 1988). Ig H chains have three or more C domains while Ig L consists of only one C domain (Schroeder & Cavacini 2010).

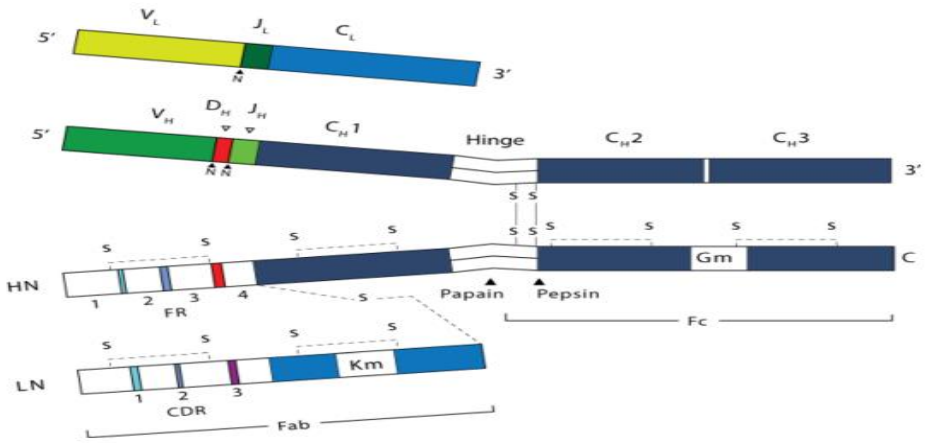


Figure 1.2. Structure of IgG showing heavy and light chains and their specific sites. (H-Heavy chain, L-Light chain, N-Amino terminus, C-Carboxy terminus, s-s-Disulfide bridge, Gm/Km-Allotype (genetic marker), V-Variable region, D-Diversity region, J-Joining region, C-Constant region, FR-Framework region, CDR-Complementary determining region) (source: Schroeder & Cavacini 2010).

During evolution, adaptive immune system played a key role in jawed vertebrates by high affinity binding to particular antigens and removal of invading pathogen. This can be done due to their characteristic feature of unlimited diversity and development of memory (Flajnik 2002). Based on the environment and antigens exposed, different species evolved various structural features in their immune system. In order to produce Ab against all foreign objects and infectious agents, the required number of genes may exceed the amount of whole genome. Therefore, strategies like V(D)J recombination and somatic hypermutation takes place in all jawed vertebrates (figure 1.3) (Tonegawa 1983; Flajnik 2002; Wang et al. 2013).

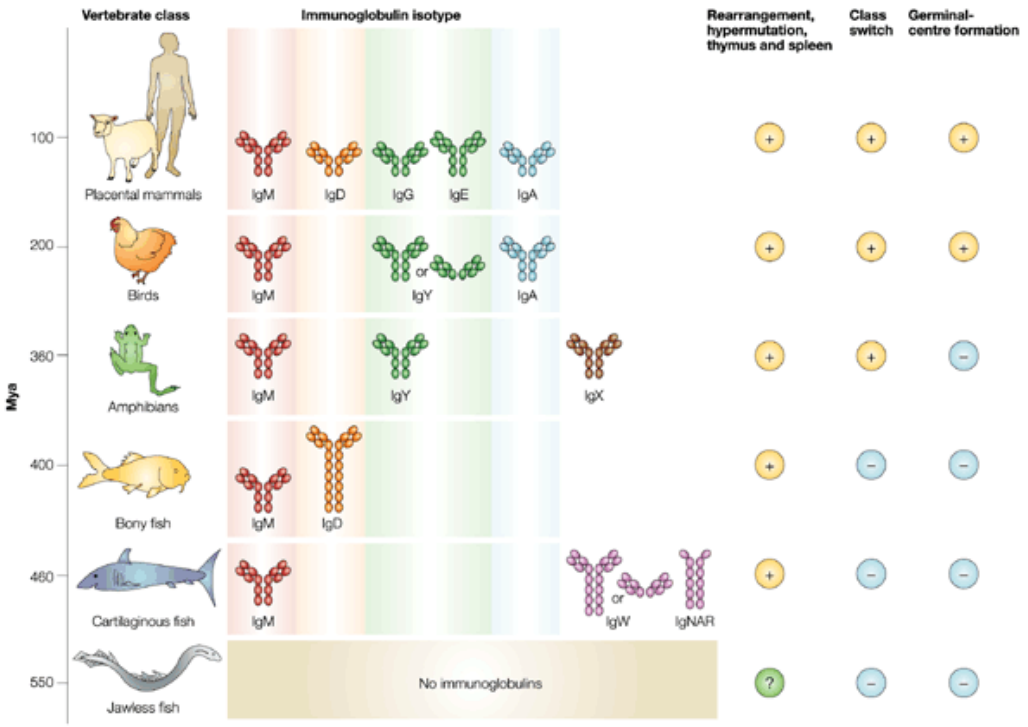


Figure 1.3. Evolution of immunoglobulins in vertebrates and development of their characteristic features (source: Flajnik 2002)

1.1.3 Shark immune system

Sharks belong to the class chondrichthyes, subclass elasmobranch. They diverged from the evolutionary common ancestor with jawed vertebrates over 500 million years ago. Cartilaginous fish are the oldest jawed species that possess acquired immune system (Flajnik 2002). However, they are devoid of bone marrow and lymph nodes (true lymphoid system) which are the major functioning units of immune system in vertebrates. But, sharks have unique immune tissues to maintain immune function, called epigonal organ and Leydig organ. It was found to be consisted with recombination activation gene (RAG) and terminal deoxynucleotidyl transferase (TdT) that is similar to

the bone marrow which is a primary lymphoid organ (Rumfelt et al. 2001; Miracle et al. 2001). In addition to the above tissues, all sharks have the thymus and the spleen (figure 1.4). Several shark species have been studied so far to determine the level of Ig expression in different organs. Spleen has higher expression of Ig in embryonic stage but none in the liver and the kidney. The Leydig and epigonal organs are found to be the major tissue which produce B-cells (Dooley & Flajnik 2006).

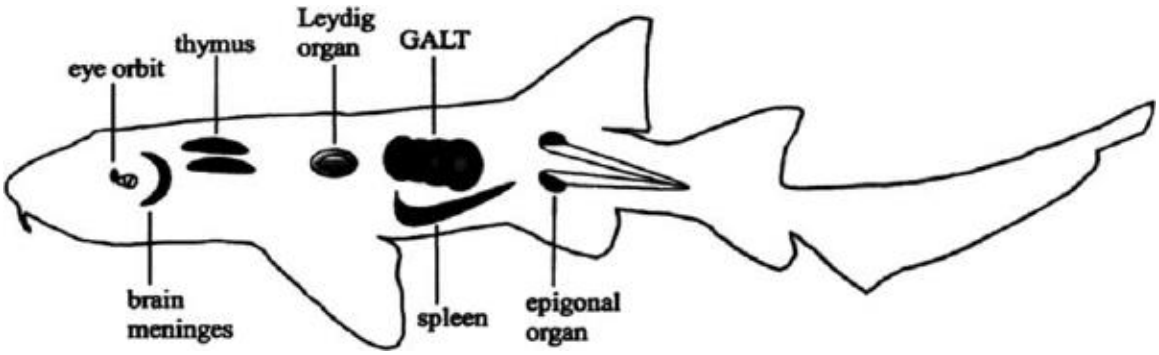


Figure 1.4. Illustration of immune tissues and organs in shark (source: Dooley & Flajnik 2006)

Not only the immune tissue but also the distribution of Ig genes is different in sharks compared to vertebrates. The Ig loci are found in cluster configuration in sharks while higher vertebrates have translocon arrangement (figure 1.5). In each cluster V, (D), J and C region exons reside and form three heavy chain Ig isotypes namely, IgM, IgW and IgNAR (novel antigen receptor) (Hinds & Litman 1986; Litman et al. 1999). Unlike vertebrate immune system, shark Igs are encoded by different clusters and there is no rearrangement among clusters and isotype switching (Dooley & Flajnik 2006).

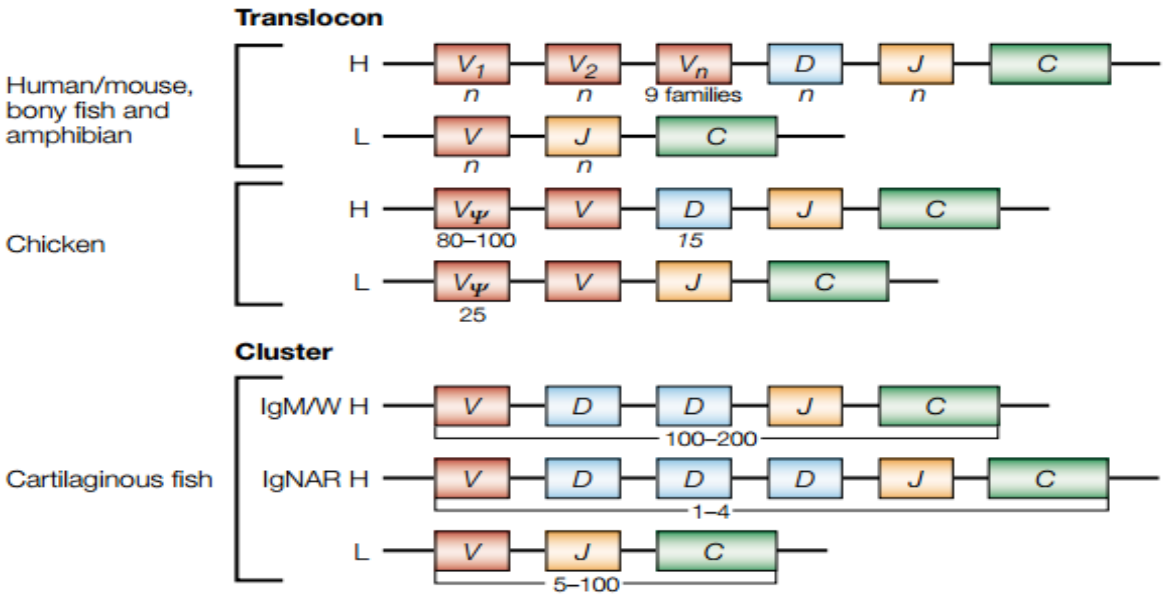


Figure 1.5. Translocon and cluster organization of immunoglobulin genes among species (source: Flajnik 2002)

Sharks represent the most primitive and fundamental features of the human immune system. Since the identification of first immunoglobulin, IgM in sharks over 50 years ago (Dooley & Flajnik 2006), many experiments have been conducted to investigate the complexity of shark immune system and their applications in the field of bioengineering.

1.1.4 Biomedical potential of IgNAR

The antibody based biotherapeutics are on the rise since recent decades. The high specificity and to avoid adverse off-site toxicity on normal tissues such as complement-dependent cytotoxicity (CDC) and antibody-dependent cytotoxicity (ADC) are challenges of bioengineering of therapeutic immunoglobulin (Presta 2008; Müller et

al. 2012). Although there are number of immunotherapeutics available to date, there are disadvantages such as, structural complications and large size of globular proteins which limits the access to cryptic epitopes (Müller et al. 2012). Therefore scientists are now concerned on the development of smaller, pH and temperature resistant (stable) antibody with higher affinity and specificity (Kubota et al. 2009).

Recent concerns on naturally occurring antibodies with single chain such as nanobodies from camelids (VHH) and IgNAR from cartilaginous fish are promising molecules for immunotherapeutics (Hamers-Casterman et al. 1993). Both antibodies are heavy chain homodimers which lacks light chain. Their smaller size (12-13 kDa), higher solubility and stability increased their potential in therapeutics. However, IgNAR variable domain (vNAR) is known as the smallest antigen binding domain available to date in the kingdom of animalia (Zielonka et al. 2015). Another significant difference between VHH and vNAR is that camelid antibody is an IgG variant while IgNAR is encoded at particular loci (cluster) (Hamers-Casterman et al. 1993). Furthermore, charged and polar amino acids in the IgNAR structure enables it to show higher solubility than other antibodies, which is an advantage to reach target epitopes (de los Rios et al. 2015). Therefore, in this study we have focused on analyzing the IgNAR structure of brownbanded bamboo shark (*Chiloscyllium punctatum*), which has not been studied to date.

IgNAR was first isolated from the serum of nurse shark (*Ginglymostoma cirratum*) and reported in 1995 by Flajnik and colleagues (Greenberg et al. 1995). It is composed of V domain with one amino terminal and C region ranging from three (transmembrane) to five (secretory) domains. Homodimerization of first and third C

domains resulted in the stability of IgNAR antibody and also consist of stabilizing motifs when engineered into mammalian antibodies (Feige et al. 2014).

IgNAR has relatively simple genetic encoding for variable region which is easy to manipulate for biotechnological and therapeutic applications. Several studies have been conducted on affinity maturation of IgNAR in elasmobranch species and phage display technologies were established to obtain IgNAR antibody (Dooley et al. 2003; Flajnik & Dooley 2009). With the help of these technologies, developments of shark recombinant IgNAR against numerous human diseases are currently studied by many research groups. Human and IgNAR Ig heavy chain variable domain amino acid identity is around 25%. Therefore several recent studies were performed humanization of vNAR domain using human serum albumin (HSA) but it showed less affinity over original antibody (Kovaleva et al. 2014). Therefore, further study on bioengineering IgNAR is a timely necessity and the identification of a suitable model organism which has not yet been investigated might reveal novel insights to the field of immunotherapeutics.

1.1.5. Monoclonal antibody synthesis and Chicken Hamster Ovary (CHO) cell expression system

Monoclonal antibodies (mAb) are derived from identical immune cell clones which originated from the unique parent cell. Their affinity is monovalent hence, bind to the same antigenic epitope. Due to its specificity, monoclonal antibody production increased the over the past decades in pharmaceutical industry. Several methods are being used to develop monoclonal antibodies including phage display, B cell culture, cell

purification and plasma cell interrogation (Wardemann et al, 2003; Huang et al, 2013; Seeber et al, 2014). Compared to 1980s, which was 10% approvals for monoclonal antibody drugs among biological products, an increase up to 30% were observed in recent years and expanding furthermore (Walsh 2014). Mammalian cell expression platforms were commonly used to reduce the immunogenicity, of which CHO cells becoming more popular. Among its advantages, firstly, the ability to achieve post translational modification making it possible to synthesize recombinant proteins having human compatible glycoforms; secondly, proved safe for host cells during clinical applications allowing it easier for approval and lastly, ability to adapt serum-free suspension culture for large scale production in bioreactors (Omasa et al. 2010).

In this study, immunization was performed using hen egg lysozyme (HEL) and developed anti-HEL antibody gene sequences based on vNAR of brownbanded bamboo shark. For the expression of the gene, mammalian expression vector (recombined pEHX1.1) (figure 1.6) and CHO cells were used.

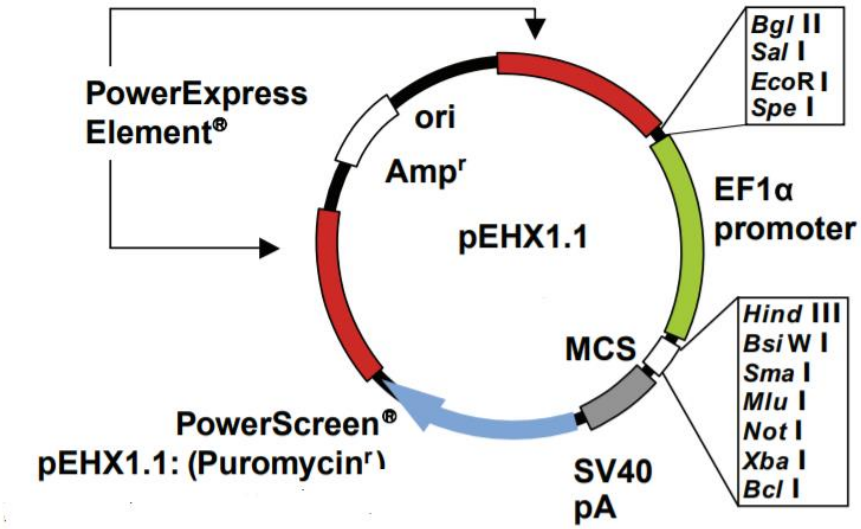


Figure 1.6. pEHX1.1 original vector map (source: www.lifescience.toyobo.co.jp)

1.2 Objectives of the study

Cartilaginous fish are the evolutionary oldest species that possess acquired immunity. Among three immunoglobulin isotypes, a heavy chain homodimer, IgNAR draws attention due to its unique structure. Smaller size (~12 kDa), higher stability and specificity to antigens as well as the ability to modify antigen binding motifs leads to ground breaking research. Currently several species of sharks have been studied for IgNAR structure and affinity, namely, nurse shark (*Ginglimostoma cirratum*) (Dooley et al. 2003), wobbegong shark (*Orectolobus maculatus*) (Nuttall et al. 2001), spiny dog-fish (*Squalus acanthias*) (Müller et al. 2012), spotted cat-shark (*Scyliorhinus canicula*) (Crouch et al. 2013) banded hound shark (*Triakis scyllium*) (Honda et al. 2010) and bamboo shark (*Chiloscyllium plagiosum*) (Zielonka et al. 2014). In each species reported variations were found in their immunoglobulin repertoire. For example, IgNAR type I was identified only in nurse shark (Greenberg et al. 1995) to date and spiny dog-fish showed high affinity IgNAR immune against human serum albumin (HAS) (Müller et al. 2012) while spotted cat-shark was unable to develop antigen specific vNAR (Crouch et al. 2013). Therefore, studying more shark species is necessary to reveal their Ig repertoire and in our experiment we chose to analyze brownbanded bamboo shark (*Chiloscyllium punctatum*) which is not yet studied by other researchers to the best of our knowledge.

Most of the immunogenetic studies were conducted to identify target specific antibodies by using phage display (Flajnik & Dooley 2009), ribosome display (Kopsidas et al. 2006) and yeast display (Zielonka et al. 2014) technologies. The development of immunoglobulin library was also performed by randomization methods and polymerase

chain reaction (PCR) methods followed by sequencing of expression clones (Diaz et al. 2002; Dooley et al. 2006). Although the use of next generation sequencing (NGS) method can reveal massive amount of information, many studies have not used it probably due to high cost. A recent study published by Feng and colleagues performed NGS of the phage displayed clones of nurse shark vNAR revealing more promising results than conventional methods (Feng et al. 2019). In this study we have performed next generation sequencing (Illumina, MiSeq) for the first time without conventional cloning methods to identify Ig repertoire and to isolate potentially highest specific vNAR gene sequence of brownbanded bamboo shark and to express using mammalian vector and CHO cell platform.

1.3 Outline of the thesis

This thesis is composed of six chapters. Chapter one is the general introduction which briefs the information regarding the immune systems of species and their changes during evolution. It also describes about structure and functions of immunoglobulins, shark immune system and the biomedical potential of IgNAR. Furthermore, it describes the application of mammalian cell line CHO on biomedical production of monoclonal antibodies.

The second chapter explains the identification and characterization of brownbanded bamboo shark IgNAR constant heavy (CH) chain domain sequences. This chapter compares the similarities and differences of *C. punctatum* CH with other shark

species vNAR studied to date. The results indicated the various types of constant domains and the diversity of brownbanded bamboo shark IgNAR.

Chapter three describes the diversity of vNAR gene sequences of *C. punctatum* due to the exposure of HEL antigen. This chapter focused on the novel method adopted to determine the variability of vNAR region in cDNA library prepared using modified Illumina MiSeq amplicon sequencing method. It reveals the dynamics of complementary determining regions (CDR) before and after antigen exposure.

The fourth chapter is about the transcriptome analysis of brownbanded bamboo shark total RNA. The objective of this study was to determine the cellular response towards antigen stimulation. Hence compared the peripheral blood leukocyte cDNA before and after HEL antigen exposure. This gives an overall picture of how the immune cells and other pathways react against antigen in brownbanded bamboo shark.

In fifth chapter, expression trial of the anti-HEL vNAR monoclonal antibody was described. The identification procedure of the highest affinity vNAR gene sequence and the novel protocol of antibody synthesis was mentioned in detail.

Final chapter reviews the results of the whole study as general discussion. Furthermore, it includes the future research potential. As future research potentials of monoclonal antibody synthesis, results of the anti-bovine serum albumin (BSA) immunization trial were presented followed by, major conclusions, challenges during this project and graphical summary.

Chapter 2

Characterization of brownbanded bamboo shark (*Chiloscyllium punctatum*) IgNAR constant domains

Some contents of this chapter were published as:

De Silva DPN, Tan E, Mizuno N, Hosoya S, Reza Md S, Watabe S, Kinoshita S and Asakawa S. Transcriptomic analysis of immunoglobulin novel antigen receptor (IgNAR) heavy chain constant domains of brownbanded bamboo shark (*Chiloscyllium punctatum*). *Fish and Shellfish Immunology*. 2018, 84:370-376

Abstract

Cartilaginous fish immunoglobulins consist of a unique antibody which is heavy chain homodimer called immunoglobulin novel antigen receptor (IgNAR). Small size, higher physicochemical stability and specificity of antigen binding domains in the variable region (vNAR) make it a potential molecule for bioengineering. IgNAR types were classified based on their variable region and so far, type I was found only in nurse shark while several other adult shark species had only type II IgNAR. Therefore, characterization of IgNAR in novel species is important not only to identify the IgNAR types but also to learn the structural stability. In this study, brownbanded bamboo shark (*Chiloscyllium punctatum*) was selected as the model species to study the IgNAR CH sequence by next generation sequencing (NGS) method. A cDNA library was synthesized from mRNA derived from peripheral blood leukocytes (PBL) and raw Illumina sequences were assembled by Paired End Read Merger (PEAR) and aligned using currently available IgNAR sequences of other cartilaginous fish species. Total of 20 million raw reads with average sequence length of 218.4 bp (35 - 301 bp) and 4.5 billion residues were obtained from Illumina sequencing. According to the transcriptome analysis, two distinct types of IgNAR were present in brownbanded bamboo shark. The C1 domain displayed 13 unique sequences that may indicate the number of IgNAR CH gene clusters. Phylogenetic analysis revealed the relationship with order Orectolobiformes, of which nurse shark (*Ginglymostoma cirratum)* and wobbegong shark (*Orectolobus maculatus*) were the closest relatives. Through this study, the knowledge on the structural features of IgNAR constant domains that may help to increase the stability could be useful in stabilizing the recombinant immunoglobulins in the future.

2.1 Introduction

Humoral immune system has shown great adaptations during the evolution of species. Development of antigen binding domains for numerous epitopes requires diversity among Ig paratopes. Therefore various species develop novel structural features in Ig repertoire depending on their natural environment (de los Rios et al. 2015). The first jawed species which challenge the antigens with acquired immunity are cartilaginous fish which consist of sharks, skates and rays (Flajnik 2002). Cartilaginous fish were diverged from jawed vertebrate ancestor around 500 million years ago (MYA), which was further divided into Holocephali (ratfish and chimera) and Elasmobranchi (rays, skates and sharks) about ~420 MYA. Along with evolution, their adaptive immune system had been developed not only with immunoglobulins, T-cell receptors and major histocompatibility complex (MHC), but also with activity of recombinase-activating gene (RAG) and somatic hypermutation (Flajnik & Kasahara 2010).

IgNAR is one of the three immunoglobulin isotypes present in cartilaginous fish. They can be isolated naturally from the shark serum in covalently bound heavy chain homodimer form. The plasma concentration of IgNAR is estimated to be around 0.1 – 1.0 mg/mL (Greenberg et al. 1995; Dooley et al. 2006). The secretary form of IgNAR consists of five constant domains while transmembrane forms have three and five constant domains in addition to their variable domain (figure 2.1). The dimerization of IgNAR heavy chain occurs via five constant domains hence both variable domains are independent from each other (Greenberg et al. 1995).

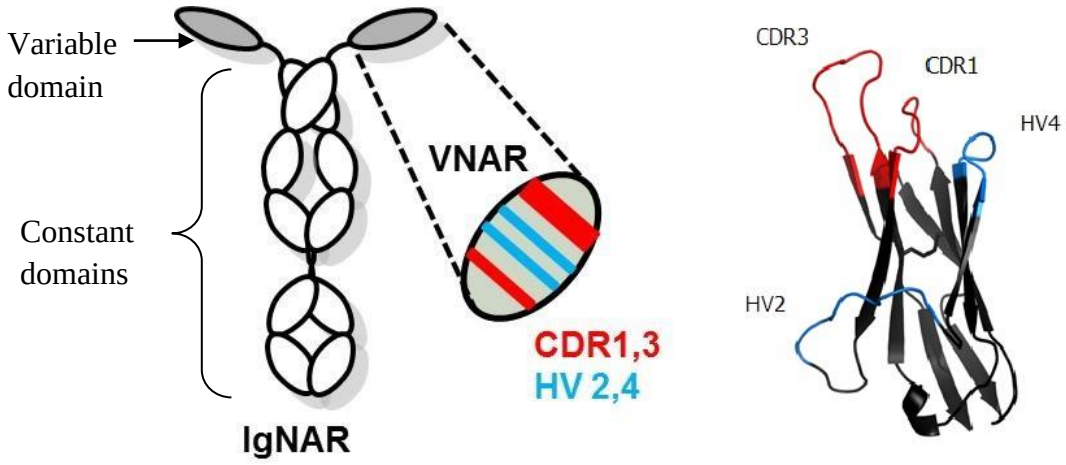


Figure 2.1. Structure of IgNAR showing variable domains (CDR1, CDR3, HV2 and HV4) and five constant domains (source: www.almacgroup.com)

In evolution, the IgNAR is believed to be derived from the cell-surface receptor and it also has homology with T-cell receptor. It is different from the other naturally available heavy chain antibody from camel known as VHH, because of its origin comes from the IgG lineage (Streltsov et al. 2005; Flajnik et al. 2011). There are three IgNAR types described based on variability of number and configuration of cystein residues in framework region and the time of its appearance in sharks during development (Rumfelt et al. 2001). IgNAR type III appears in early development of sharks and believed to have defense against infections at early stages. In nurse shark both IgNAR type I and II levels increase upon exposure to antigens in mature sharks (Diaz et al. 1998). The difference of cysteine amino acid position among three immunoglobulin types is shown in figure 2.2. Type I IgNAR is only found in nurse shark to date and all the other shark species examined to date possess type II and its variations. Therefore, investigation of IgNAR sequence of other shark species is very important in order to identify more IgNAR types.

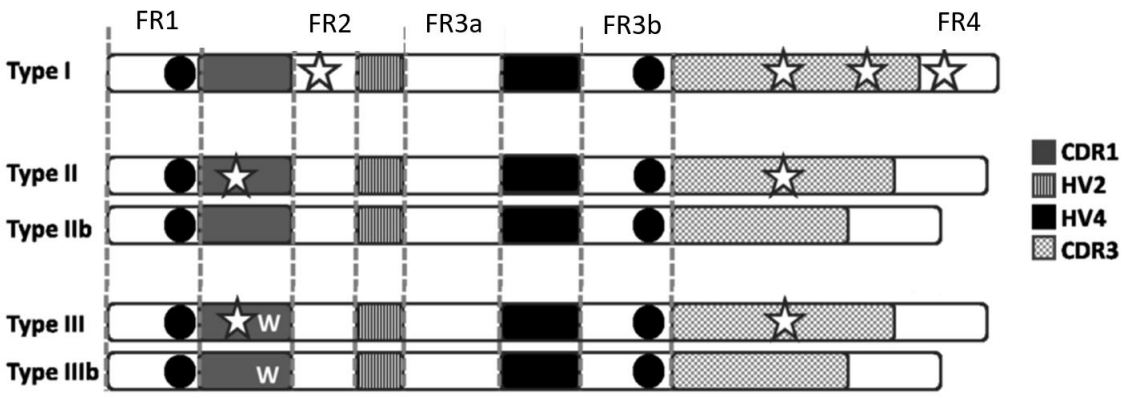


Figure 2.2. Classification of IgNAR based on the non-canonical cysteine residue in variable region. Solid black filler circles - canonical cysteine residues; stars - non-canonical cysteine residues; FR – framework (source: Barelle & Porter 2015).

Nurse shark IgNAR has been studied extensively over the past decades and several other studies were conducted on various shark species. Zielonka and colleagues (2014) experimented on *in vitro* development of semi-synthetic vNAR molecule using non-immunized whitespotted bamboo shark (*Chiloscyllium plagiosum*) which is the first study on using bamboo shark as a model species (Zielonka et al. 2014). Brownbanded bamboo shark (*Chiloscyllium punctatum*) which is in the same taxonomic order with nurse shark and wobbegong shark has not been investigated for their immune repertoire to date. The presence of species variation of IgNAR types and their sequence structure arose the importance of investigating new shark species in detail. classification of the IgNAR clusters based on the variable region alone might not be sufficient mainly due to the higher variability in the CDR regions. Hence, this study proposes the possibility of using the C1 domain sequence to identify IgNAR clusters due to its uniqueness compared to the other four C domains which share sequences similarity with IgW (Marchalonis et al. 1998). Furthermore, this study aimed to characterize the complete IgNAR heavy chain

constant domain of brownbanded bamboo sharks (*Chiloscyllium punctatum*) and to determine its evolutionary relationship with other species.

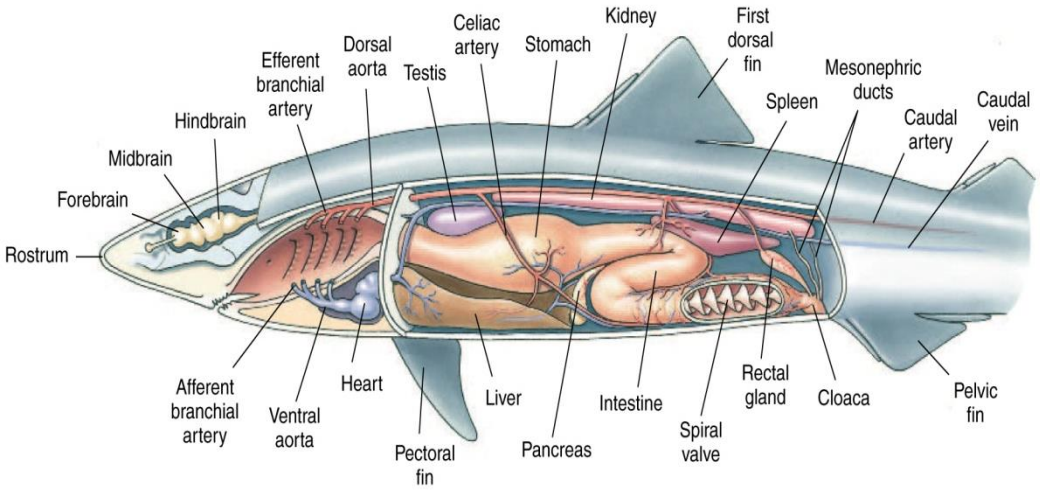
The application of next generation sequencing methods have not been done by many researchers to date on IgNAR except in *Triakis scyllium* IgNAR analysis study that used Genome Sequencer FLX (GS FLX) in their study (Honda et al. 2010) and by Feng et al. for naïve nurse shark phage display library preparation (Feng et al. 2019). The amplification of IgNAR for the first time of a particular species were performed in almost all the previous studies by primer designing to reference sequence of another species or cloning. These methods may lead to biasness and restrict the volume of data which mask the actual scenario. Therefore in this study, cDNA library was synthesized by the total RNA extracted from peripheral blood leukocytes (PBL) through Illumina (MiSeq) sequencing technique. Transcriptomic analysis using various bioinformatic tools revealed novel information and enhanced the IgNAR gene sequence database which will be described in results and discussion section of this chapter.

2.2 Methods

2.2.1 Sample collection

Adult brownbanded bamboo sharks (*Chiloscyllium punctatum*) raised at NIFREL aquarium (Osaka, Japan) were transferred to Fisheries Laboratory, Graduate School of Agricultural and Life Sciences, The University of Tokyo (Shizuoka, Japan). All the experimental procedures were performed with live sharks by following the institutional animal ethics guidelines. Five sharks in total were used in this experiment. Prior to sampling, sharks were anaesthetized by spraying MS-222 (~250 µg / 500 ml sea water) via gill slits and oral routes (Gilbert & Wood 1957). The caudal vein was punctured by using 18 gauge needle attached to a 25 cm blood collection tube (Terumo, USA) and the pressure applied by 10 ml sterile syringe (Terumo, Japan) as shown in figure 2.3.

a.



b.

Figure 2.3. a. Internal anatomy of shark (source: www.kln.ac.lk/science/depts/zoology)

b. Collection of blood from the caudal vein of anaesthetized brownbanded bamboo shark

Whole blood around 8 ml were collected and transferred to 15 ml falcon tubes (Fisher Scientific, USA) containing 500 U of heparin (Japan Pharmacopea, Japan) and mixed by shaking to prevent coagulation. In order to extract the peripheral blood leukocytes (PBL), centrifugation method stated by Smith and colleagues (2004) was used with slight modifications (Smith et al. 2004). The steps of isolation of PBL were as follows: Whole blood samples were centrifuged at 600 g for 15 minutes. Then plasma sample about 5 ml were collected to a separate tube and PBL (Buffy coat layer) were transferred using 1000 µl micropipette in to another 15 ml tube. The separated PBLs were diluted using 5 ml of elasmobranch-phosphate buffered saline (E-PBS), which was made by mixing NaCl 2.63 g with NaH_2PO_4 0.12 g at pH 7.4 in 100 ml of sterile water (Smith et al. 2004). The E-PBS and PBL mixture was centrifuged at 1200 g for 10 minutes and removed the supernatant followed by adding E-PBS 5 ml into the pellet. The centrifugation and removal of supernatant was performed again and purified PBL were

suspended in RNAlater solution (Ambion, USA) and stored at -20°C until further total RNA extraction.

2.2.2. Total RNA extraction and sample processing for Illumina MiSeq sequencing

RNeasy mini kit (Qiagen, Germany) was used to extract total RNA following the manufacturers' instructions. In brief, PBL samples stored in 1.5 ml tubes were centrifuged at 6000 g (himacCF15R, Hitachi, Japan) and discarded the supernatant containing RNA later. Buffer RLT 350 µl were added and cells were disrupted by using homogenization (Precellys24, Bertin Technologies, France). Same volume (350 µl) of 70 % ethanol were added to the lysate and transferred to RNeasy spin column placed in 2 ml collection tubes followed by centrifugation at 13000 rpm for 30 seconds. The contents of the membrane were filtered through buffer RW1 and buffer RPE according to the protocol and finally nuclease free water 40 µl were added and spun down to collect the total RNA. The extracted RNA was quantified by using Qubit RNA Assay kit (Invitrogen, Carlsbad, CA) and quality was tested by Agilant 2200 tapestation system (Agilent Technologies, USA). Based on the RNA quality and quantity, the best shark total RNA sample was selected for cDNA library synthesis.

TruSeq stranded mRNA and total RNA sample preparation kit (Illumina, California, USA) was used to synthesize cDNA library following the Low Sample protocol under the manufacturers' guidelines with minor alterations. In brief, sample preparation workflow begun with the preparation of pooled 1.21 µg of total RNA to purify and fragment mRNA. First strand cDNA was synthesized by adding SuperScript II

reverse transcriptase enzyme, followed by second strand cDNA synthesis. Adenylation of 3'ends were done to prevent ligation among cDNA strands during adapter ligation. Adapter sequence GCCAAT was ligated and amplified by polymerase chain reaction (PCR) using thermal cycler (Veriti, Applied Biosystems, USA). Final amplified product of over 500 bp in length was gel extracted and purified for Illumina sequencing. Quality of cDNA was verified by Agilant 2200 tapestation and up to four nano mols (4 nmol) of cDNA were sequenced by MiSeq Sequencing System (Illumina, USA).

2.2.3 Sequence analysis using bioinformatics

Raw Illumina sequences were analyzed by using number of bioinformatic softwares. Raw sequence quality was evaluated by FastQC **(**v 0.11.3**)** **(**www.bioinformatics.babraham.ac.uk) and Trimmomatic (version 0.32) was used for adaptor trimming (Bolger et al. 2014). Sequence assembly was performed in paired ends by Paired-End ReAD mergeR (PEAR) (version 0.9.6) (Zhang et al. 2014) and Trinity software (Grabherr et al. 2011) using the default settings.. In order to identify the IgNAR sequence, a local database was created first by using a set of currently published cartilaginous fish IgNAR sequences obtained from National Center for Biotechnology Information (NCBI) nucleotide database. The reference sequences and their accession numbers are listed in table 2.1. The local alignment of assembled sequences was performed and the best matched sequences were extracted for transcriptome analysis of IgNAR CH region. Sequence editing was performed by using BioEdit Sequence Alignment Editor (version 7.2.6.1) (Hall, 1999).

Table 2.1. Cartilaginous fish IgNAR sequence NCBI accession numbers used in the local database to identify *C. punctatum* IgNAR

Species name	General name	NCBI accession	Reference
Ginglymostoma cirratum	Nurse shark	AAB42621.2	Greenburg et al., 1995
Orectolobus maculatus	wobbong shark	ABB83616.1	Simons et al., 2006
Scyliorhinus canicula	small-spotted catshark	AGG53259.1	Crouch et al., 2013
Triakis scyllium	banded houndshark	BAJ20185.1	Honda et al., 2010
Squalus acanthias	spiny dogfish	AES92985.1	Smith et al., 2012
Rhinobatus productus	Shovelnose guitarfish	AAT02204.1	Rumfelt et al., 2004

The nurse shark IgNAR heavy chain constant region amino acid sequence was the reference for characterization of the brownbanded bamboo shark IgNAR constant domains. The transcriptome sequences from the PEAR and Trinity assembly software were compared, and nucleotide variations were determined. Consensus sequences were created by assembling the contigs using the CLC Genomics Workbench version 8.0.3 (www.qiagenbioinformatics.com). The ClustalW multiple alignment program was used to align the constant domains of different species, and gene editing was performed by the BioEdit version 7.2.6.1 software (Hall 1999). The phylogenetic tree was created by the neighbor-joining method (Saitou & Nei 1987) with 1000 bootstrap replicates (Felsenstein 1985) using the MEGA v.7 software (Kumar et al. 2016) and the evolutionary distances were calculated by the Poisson correction method (Zuckerkandal & Pauling 1965).

2.3 Results

2.3.1. Amplification of brownbanded bamboo shark cDNA for next generation sequencing

Total RNA 44.4 ng / µl from the PBL with 7.2 RNA Integrity Number (RIN) (figure 2.4) was obtained from healthy adult shark. The final concentration of amplified cDNA (TruSeq Stranded RNA protocol) was 194 pmol/l and after amplification by PCR around 4 nmol were used for Illumina sequencing.

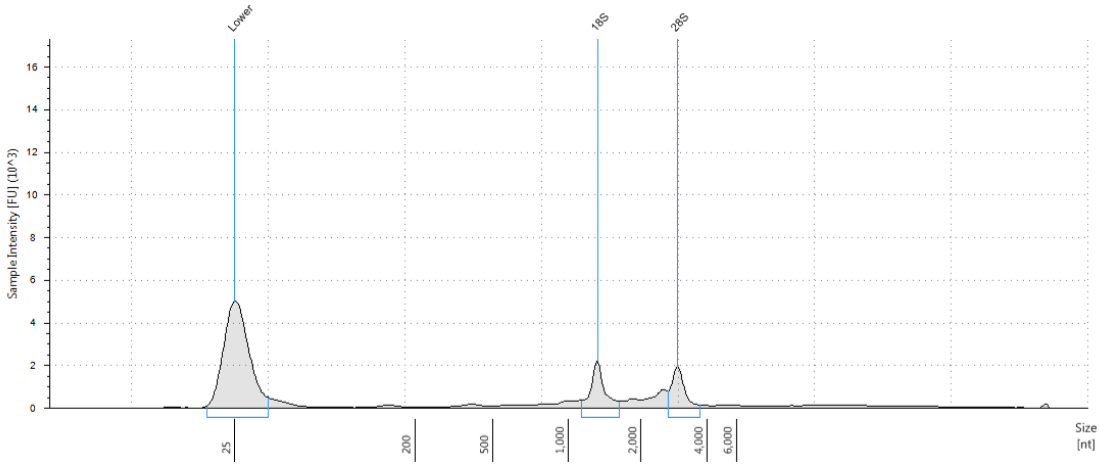


Figure 2.4. Quality of total RNA extracted from brownbanded bamboo shark PBL analyzed by Agilant tapestation 2200.

2.3.2 Brownbanded bamboo shark cDNA library

The next generation sequencing of brownbanded bamboo shark total RNA from peripheral blood leukocytes resulted in total of nine giga base (9 GB) of residues from 20 million sequences with the average length 218.41 base pairs (35 bp - 301 bp), consisted of 43% GC. Per base sequence quality results from FastQC indicates the high quality score (good quality calls) for most of the bases and gradually degraded towards

the end of the run progressed which is a normal fact of most sequencing platforms as shown in figure 2.5 (www.bioinformatics.babraham.ac.uk). The large set of forward and reverse raw sequences was reduced up to two million (48.2 %) contigs after PEAR assembly with average length of 244.12 bp (50 bp - 592 bp).

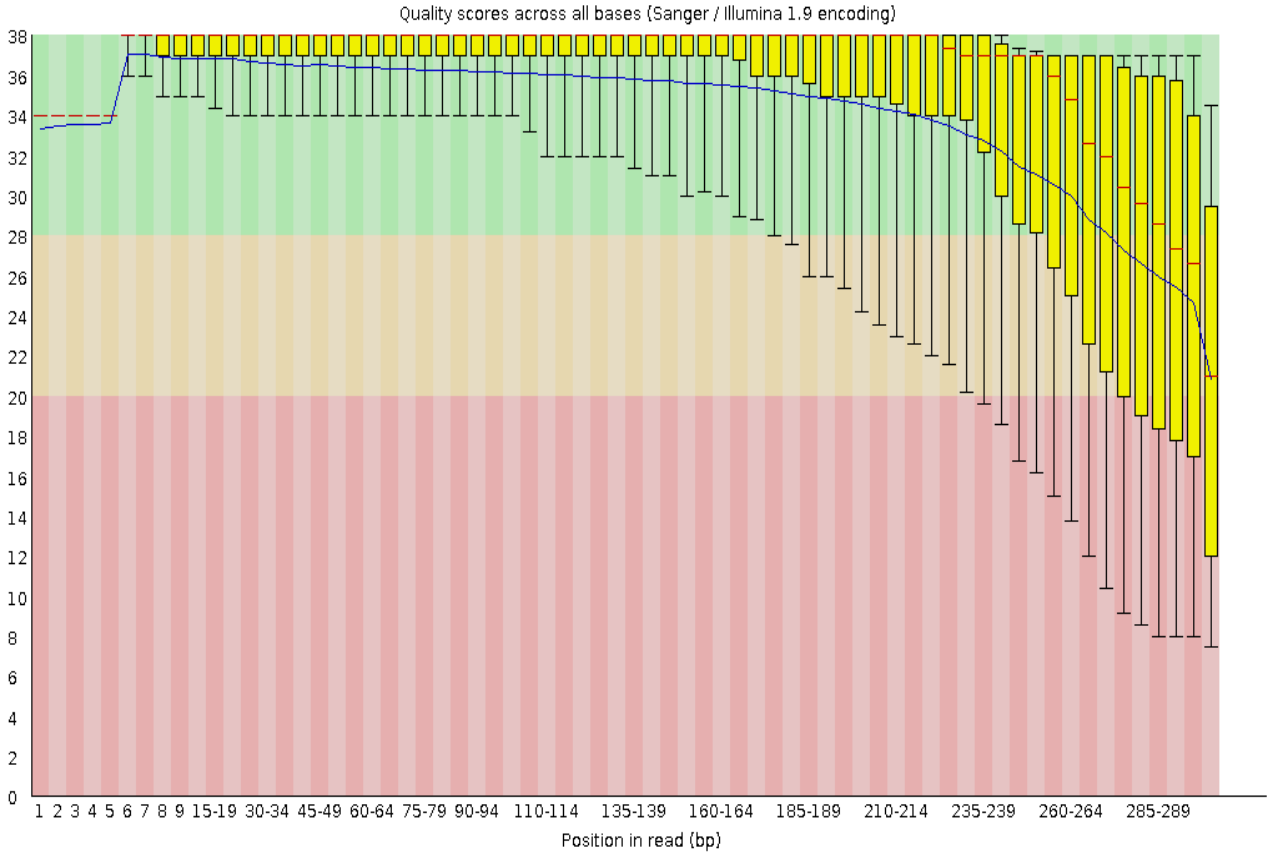


Figure 2.5. FastQC per base quality scores of brownbanded bamboo shark raw sequence (forward strands) obtained from Illumina MiSeq.

The sequence assembly by Trinity formed 0.3 million contigs while the PEAR assembly merged up to 19 million Illumina paired-end reads. The average sequence length obtained by the Trinity software was much higher than using the PEAR assembler although the number of contigs was less in the Trinity assembly as displayed in Table 2.2.

Table 2.2. Summary of the brownbanded bamboo shark transcriptome library indicating the sequence count and lengths before and after assembly by the Trinity or PEAR software

	No of sequences	Total residue count	Minimum sequence length	Maximum sequence length	Average sequence length
Raw - Forward	20,599,176	4,499,047,447	35	301	218.41
Raw - Reverse	20,599,176	4,528,135,612	35	301	219.82
PEAR	19,890,817	4,855,711,172	50	592	244.12
Trinity	303,547	209,162,459	201	25,719	689.06

2.3.3. Characterization of the IgNAR constant domains

As per the alignment, 154 cDNA sequences of the PEAR-assembled library were aligned and mapped to complete IgNAR constant domains (secretory form) using the nurse shark (NCBI accession U18721) as the reference sequence. As a result, two representatives secretory IgNAR CH chain nucleotide sequences were obtained composed of 1,758 base pairs coding 586 amino acids from the brownbanded bamboo shark (figure A1.1 in appendix). The Trinity-assembled IgNAR CH contigs were compared with the PEAR-merged sequences. A higher homology was revealed and confirmed by the ClustalW sequence identity matrix to be greater than 0.9. Based on the transcriptome sequence of C1, the brownbanded bamboo shark IgNAR was classified mainly into two types, i.e., the *C. punctatum* IgNAR-1 CH type and the IgNAR-2 CH type. A total of seven major polymorphic sites with minor variation frequency above 20 % were distributed in the C1 domains of both IgNAR types with six nonsynonymous and one synonymous mutation (table 2.3 and figure 2.6).

Table 2.3. Minor variation frequencies and respective amino acid changes in IgNAR constant heavy chain C1 domain

Mutation*	Minor variation frequency %[#]	Polymorphism
c. 57 A>C (p. Glu19Asp)	32.2	Nonsynonymous
c. 73 A>G (p. Lys25Gly)	31.5	Nonsynonymous
c. 74 A>G (p. Lys25Gly)	31.5	Nonsynonymous
c. 180 G>T (p. Ser60Ser)	36.6	Synonymous
c. 181 C>T (p. Pro61Ser)	41.5	Nonsynonymous
c. 238 A>C (p. Lys80Gln)	20.8	Nonsynonymous
c. 305 A>G (p.Lys102Arg)	42.9	Nonsynonymous

*mutation nomenclature based on Ogino et al., 2007.

[#]minor variation frequency was calculated by the percentage value of the proportion of minor variations expressed per base site in C1 domain

a.

b.

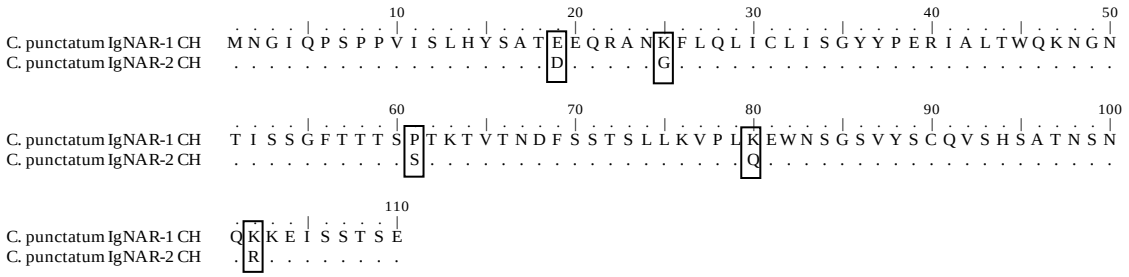


Figure 2.6. Brownbanded bamboo shark IgNAR heavy chain C1 domain sequences indicating the presence of two IgNAR types: **a)** nucleotide sequence and **b)** amino acid sequence alignments indicating the polymorphic sites. Dots represent the identical residues, and rectangular boxes demarcate the polymorphic sites.

Based on the distribution of single nucleotide polymorphisms (SNPs), 13 unique IgNAR C1 gene sequences were identified (figure A1.2 in appendix). In the bamboo shark IgNAR C region (C1–C5), 199 SNPs were found that resemble 12.6% of the total IgNAR CH base positions. From the 154 PEAR-merged contigs analyzed in this study, 3.58% of polymorphisms were found in the C3 domain followed in decreasing order by C4 (2.73%), C2 (2.33%), C1 (1.65%), and C5 (0.8%) along the full length of the secretory IgNAR. Variable nucleotides are listed in the appendix as table A1.1.

The average molecular weight of the entire IgNAR constant region of the brownbanded bamboo shark was 65.3 kDa. Based on the percentage of the total molecular weight, it is primarily composed of serine (11.58%), threonine (9.03%), and valine (8.01%). The least abundant amino acid in the brownbanded bamboo shark IgNAR constant domain is histidine (1.02%). The amino acid composition of the bamboo shark IgNAR was listed in the appendix table A1.2.

2.3.4. Phylogenetic analysis of the *C. punctatum* IgNAR constant region

The brownbanded bamboo shark IgNAR heavy chain constant region amino acid sequence compared with the NCBI blastp revealed a 77% identity to the nurse shark and was more than 57% comparable to other cartilaginous fish (table 2.4). The phylogenetic analysis revealed the evolutionary relationship of the brownbanded bamboo shark with the nurse shark and the wobbegong shark which belong to the same taxonomic order (Orectolobiformes). Based on the phylogenetic tree, the IgNAR constant region of the brownbanded bamboo shark and the nurse shark shared a common ancestor with the wobbegong shark. The phylogenetic tree and the protein sequence alignment were shown in figures 2.7 and 2.8, respectively.

Table 2.4. Comparison of the bamboo shark IgNAR constant heavy chain protein sequence using the NCBI blastp with published cartilaginous species IgNAR sequences

Species name	NCBI Accession number	Max. score	Total score	Query cover	E value	Identity
Nurse shark *(G. cirratum)*	AAB42621.2	938	996	99%	0.0	77%
Spotted wobbegong *(O. maculatus)*	ABB83616.1	869	1034	89%	0.0	79%
Spiny dogfish *(S. acanthias)*	AES92985.1	743	933	89%	0.0	67%
Banded houndshark *(T. scyllium)*	BAJ20188.1	705	1029	89%	0.0	65%
Small-spotted catshark *(S. canicular)*	AGG53259.1	655	952	89%	0.0	61%
Shovelnose guitarfish *(R. productus)*	AAT02204.1	612	668	89%	0.0	57%

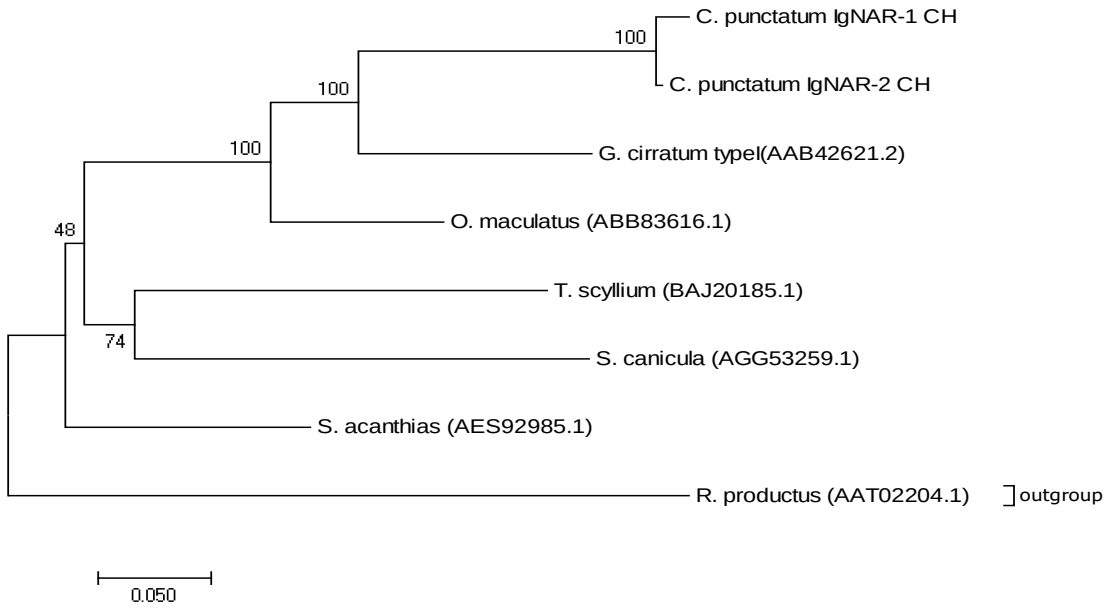


Figure 2.7. Phylogenetic analysis of the IgNAR heavy chain constant region complete sequence of the brownbanded bamboo shark (IgNAR-1 CH and IgNAR-2 CH) compared to other cartilaginous fish species. *Rhinobatos productus* is a ray used as an outgroup. All species names and their accession numbers are indicated at their respective nodes. The genetic divergence is indicated by the branch length and scaled to 0.05. The tree is drawn to scale, with branch lengths in the same units as the evolutionary distances used to infer the phylogenetic tree.

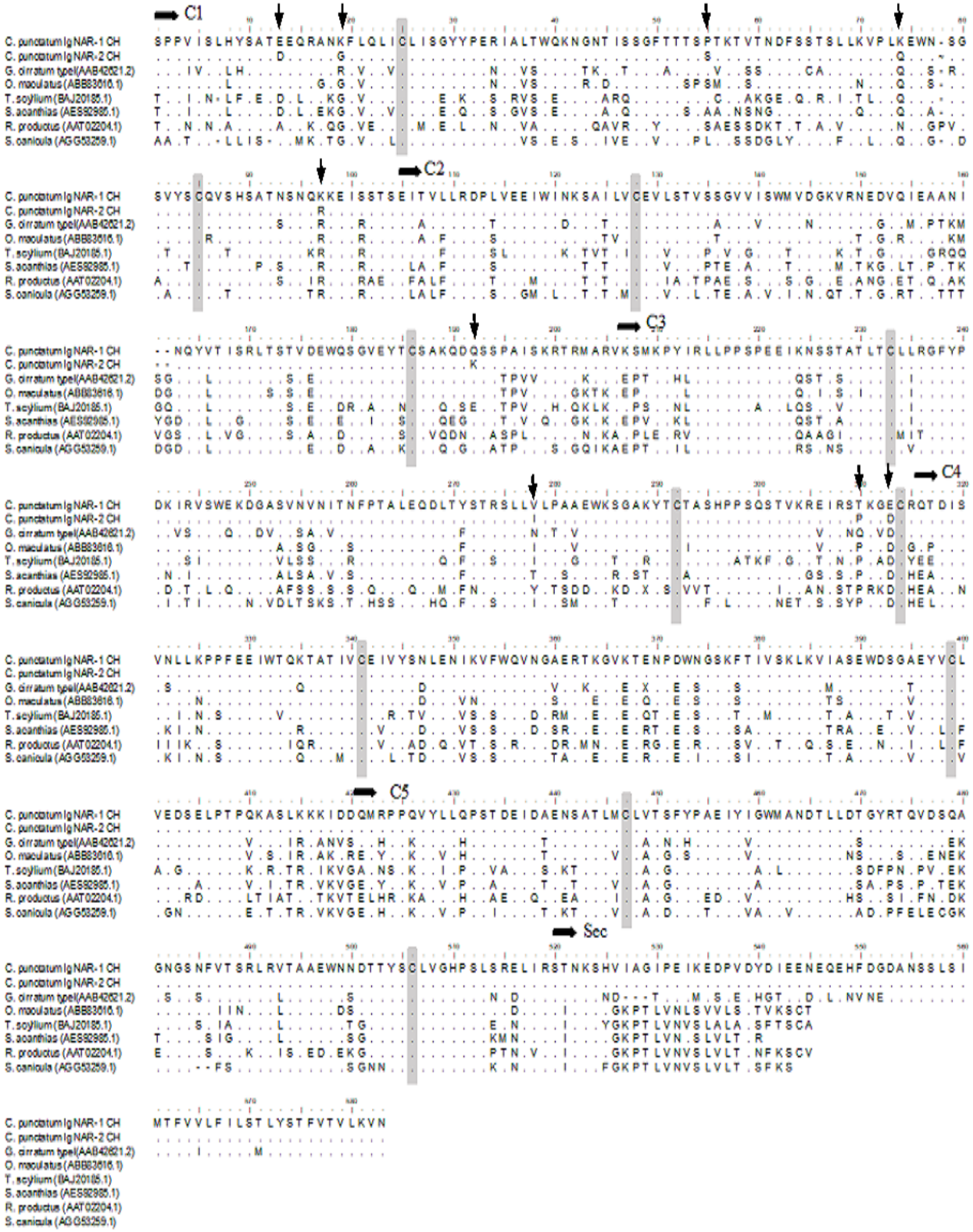

C1
C2
C3
C4
C5
Sec
C. punctatum Ig NAR-1 CH
C. punctatum Ig NAR-2 CH
G. cirratum type II (AAB42621.2)
O. maculatus (ABB83616.1)
T. scyllium (BAJ20185.1)
S. acanthias (AES92985.1)
R. productus (AAT02204.1)
S. caniculа (AGG53259.1)

Figure 2.8. The brownbanded bamboo shark IgNAR (*C. punctatum* IgNAR-1 CH and *C. punctatum* IgNAR-2 CH) constant region amino acid alignment compared to other cartilaginous fish species. NCBI accession numbers and species names are available at the left of the sequences. The black horizontal arrow marks the beginning of each constant and secretory domain (C1 to C5 and Sec). Gray areas indicate the conserved cysteine residues. Amino acids identical to the IgNAR-1 CH are marked with black dots and dashes represent gaps. Vertical black arrows indicate the nonsynonymous mutations in the IgNAR-2.

2.3.5. Compatibility of the *C. punctatum* IgNAR constant region for immune modifications

Though IgNAR sequences are dissimilar to human immunoglobulin domains, the essential elements in the immunoglobulin fold are conserved. Like the analysis by Feige et al. 2014, in this study, the brownbanded bamboo shark IgNAR constant domains (C1 to C5) were aligned with other cartilaginous fish CH domains and human IgG domains (figure 2.9).

Similar to the previous study on nurse sharks, brownbanded bamboo shark IgNAR CH also possessed a tryptophan molecule which helps hydrophobic residues to form a tight core around the disulfide bridge (Feige et al. 2014). Furthermore, molecules such as cysteine, tyrosine, phenylalanine and tryptophan along the C1-C5 domains in brownbanded bamboo shark IgNAR marked the significance of evolutionary conservation and indicates the stability in nature which in turn can be used to modify human immunoglobulins. Among the differences with the human IgG domains, the

IgNAR lacked proline residues which is essential to the mammalian immunoglobulin folding (Zuckerkandl & Pauling, 1965; Feige et al. 2014). However, in IgNAR constant domains, presence of hydrophobic core which surrounds the disulfide bridge and formation of helix interacting with aromatic residues lead to the stability and flexibility which lacks in human IgG.

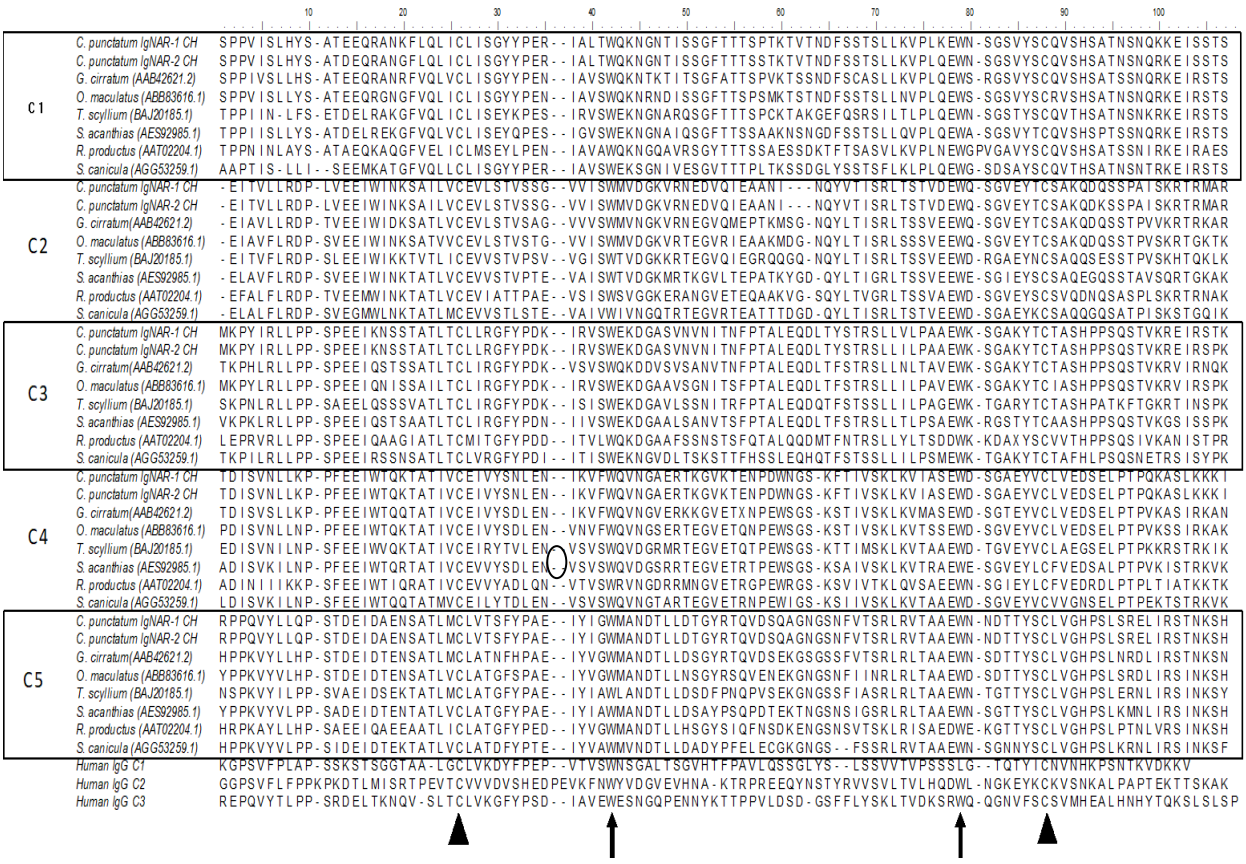


Figure 2.9. Amino acid sequence alignment of *C. punctatum* and other cartilaginous fish CH domains to Human IgG constant domains. An arrowhead marks conserved cysteine residues, and a thin arrow indicates conserved tryptophan residues. The open circle in the human IgG sequence marks the proline residue which is lacking in the IgNAR constant domains of all species.

2.4. Discussion

IgNAR is a novel form of heavy chain homodimer which has not been identified in brownbanded bamboo shark to date. The closest species studied to date was by Zielonka and colleagues (2014 and 2015) on white spotted bamboo shark (*Chiloscyllium plagiosum*). They performed *in vitro* affinity maturation of vNAR antigen binding sites against several antigens such as epithelial cell adhesive molecule (EpCAM) and Ephrin type-A receptor 2 (*EphA*2) by yeast surface display (YSD) method. Their findings revealed the presence of high affinity binders (Zielonka et al. 2014) and autonomous paratopes (Zielonka et al. 2015) in vNAR which is a promising result for future studies on bamboo sharks.

In this study, IgNAR sequences were obtained through Illumina sequencing. Illumina, a next generation sequencing platform is currently used in multiple aspects of genomic research including immune repertoire sequencing. Immunoglobulins produce vast number of receptors due to their rearrangement events. Therefore NGS play a key role in sequencing unknown, diverse set of sequences (Benichou et al. 2012), hence used in this study to synthesize brownbanded bamboo shark cDNA library.

Lack of reference sequence to map the *de novo* assembly was a critical point in this study. However preparation of local IgNAR sequence database from already published sequences of other shark species were aligned to obtain the best hit sequences of brownbanded bamboo shark IgNAR.

Since the first record of IgNAR by Flajnik and co-workers in 1995, several cartilaginous species have been studied by different research groups. Nurse shark

(*Ginglymostoma cirratum*) was the most studied species of all which has mainly two types of IgNARs in adults. The earliest classification was based on three criteria such as, presence of a cysteine in strand C (later FR2 domain) of variable region and cysteine in J region (later FR-4 domain) in type I, Cysteine in CDR1 domain named as type II and heterogeneous CDR3 consisting 1-4 cysteines in both types (Greenberg et al. 1995). Hence, current categorization of the IgNAR is primarily based on the location of the non-canonical cysteine residue in the variable region. Because the type I IgNAR is found only in nurse sharks, while all other species express the IgNAR type II and its variants (Flajnik 2002), another means of categorization is necessary. Therefore, this research proposed the possibility of a novel classification of the IgNAR, based on the IgNAR C1 nucleotide variations. The polymorphisms in the C1 domain were considered due to the uniqueness in the IgNAR constant region compared to the other four domains which are derived from IgW (Flajnik & Rumfelt 2000; Rumfelt et al. 2001). The IgNAR C1 domain also plays a significant role in structural alterations that increase affinity against a specific antigen (Fiege et al. 2014). From the findings of this study, the brownbanded bamboo shark IgNAR-1 CH and IgNAR-2 CH sequences can be used to compare novel IgNAR types in future studies on cartilaginous fish.

In this study, due to the complexity and the number of sequences, two software programs with different assembling systems were used to obtain a more accurate analysis. The Trinity assembly provided longer reads covering more than half of the entire IgNAR constant region while the PEAR assembly merged the paired ends of the Illumina reads, giving rise to several short contigs with higher accuracy. Both analytical methods

covered the entire IgNAR constant region transcriptome sequence and proved to be suitable to analyze NGS data.

Cartilaginous fish diverged from their jawed vertebrate ancestor approximately 450 million years ago and further divided into several lineages, particularly forming the superorder Galeomorphi (approximately 220 million years ago) and bringing the brownbanded bamboo shark (*C. punctatum*) into existence (Zuckerkandl & Pauling 1965). During the evolution of sharks, the Squalea and Glaleomorphii lineages experienced cluster expansion or contraction, resulting in a different number of clusters. In spiny dogfish, it appears that immunoglobulin clusters are arranged tandemly, indicating the possibility of duplication and expansion of the primordial cluster during meiosis due to unequal cross-over events (Smith et al. 2012). Gene conversion may homogenize the immunoglobulin clusters to maintain high sequence identity, increasing the chance of more expansions or contractions. Similar observations have been made in the brownbanded bamboo shark IgNAR C1 domain transcriptome sequence, resulting in nucleotide variations in 13 sequences, which might indicate that the brownbanded bamboo shark may contain up to 13 independent IgNAR gene clusters. Interestingly, all five nonsynonymous mutations in the brownbanded bamboo shark IgNAR C1 domain were also found in other shark species which might have been retained during evolution. Because the IgNAR lacks the canonical hinge region, the C1 dimerization interface produces a wide angle, facilitating the variable domain to bind multiple epitopes, which is another reason for the increased complexity of the IgNAR C1 domain compared to the other constant domains. However, the overall number of nucleotide variations per domain is highest in the C3 domain in bamboo sharks possibly due to the formation of a narrow

stalk by C3 dimerization which supports the flexibility of the IgNAR using a disulfide bridge between the C3 and C4 domains (Zielonka et al. 2015).

Molecular analysis and characterization of shark antibodies are essential not only due to its evolutionary history but also to the study of their existence in harsh conditions such as high urea concentrations in blood (Dooley & Flajnik 2006). The brownbanded bamboo shark, similarly to other sharks, consisted of 11 conserved cysteine residues in its IgNAR constant domains. Furthermore, the alignment of the IgNAR constant domains with other cartilaginous fish and human IgG revealed that brownbanded bamboo sharks also possess conserved tryptophan and cysteine residues that are involved in the formation of disulfide bonds and structural folds (Fiege et al. 2014).

The composition of amino acids in the IgNAR constant domains is significant for the stability of the antibody because it does not have support from the light chain domains (Fiege et al. 2014). The structural elements in the IgNAR constant domain could also influence the folding pathway of the antibody which increases the stability. This attribute makes it a suitable candidate as a therapeutic antibody, mainly due to the ability to transplant the stabilizing structural motif into other domains. Furthermore, the domains in the IgNAR are homologous with IgW, which is one of the oldest isotypes among immunoglobulins (Dooley & Flajnik 2006) and IgNAR is one of the immunoglobulin isotypes which is responsible to antigen specific response upon immunization (Dooley & Flajnik 2005).

This study was conducted using transcriptomic sequences obtained from next-generation sequencing (Illumina, MiSeq) and the results were compared by using two assembling methods. Both methods provided nearly similar outcomes, although the

Trinity software provided longer sequence reads compared to the PEAR assembly procedure. This study is the first record of the detailed characterization of the brownbanded bamboo shark IgNAR constant domain to the best of our knowledge. Coding sequences of the *C. punctatum* IgNAR constant region were deposited at the NCBI GenBank. Accession numbers for IgNAR-1-CH and IgNAR-2-CH are MH813472 and MH813473 respectively.

Chapter 3

Diversity of brownbanded bamboo shark (*Chiloscyllium punctatum*) vNAR in response to antigen stimulation

"As the contents of this

chapter are anticipated to be published in a paper in a scholarly

journal, they cannot be published online. The paper is scheduled to be

published within 05 years."

Chapter 4

Transcriptome analysis provide insights to gene expression profile of lymphoid cells in brownbanded bamboo shark (*Chiloscyllium punctatum*)

"As the contents of this chapter are anticipated to be published in a paper in a scholarly journal, they cannot be published online. The paper is scheduled to be published within 05 years."

Chapter 5

Selection of the antigen specific vNAR candidate gene by expression analysis

"As the contents of this chapter are anticipated to be published in a paper in a scholarly journal, they cannot be published online. The paper is scheduled to be published within 05 years."

Chapter 6

General discussion

6.1. General discussion

Cartilaginous fish as the evolutionary oldest jawed vertebrates which possess acquired immunity, have the greatest potential of having the basic immunoglobulins. IgM is one of the antibodies which are expressed in all vertebrates including humans (Dooley & Flajnik 2006). Another antibody which can be found in cartilaginous fish is IgW. It has many forms based on the length of constant domains. This immunoglobulin is considered as the IgD antibody in other animals (de los Rios et al. 2015). The other immunoglobulin studied in details of present study was IgNAR. Which devoid of light chains and a heavy chain homodimer (Greenberg et al. 1995).

Since the first record of IgNAR in 1995, number of studies carried out to investigate the immunoglobulin structure, function and affinity maturation. Though the exact function of IgNAR was not clear to date, its role in adaptive immunity and memory was identified by various researchers (Dooley & Flajnik 2005). IgNAR draws attention for various applications as an immunotherapeutic and an immunodiagnostic molecule hence this study investigated the IgNAR sequence using transcriptomics to determine the possibility of introducing brownbanded bamboo shark as an animal model. Nurse shark is the widely studied species in terms of IgNAR and brownbanded bamboo shark belongs to the same taxanomic order with nurse shark. Another reason for choosing this model was that it is easy to rare brownbanded bamboo shark in captivity and their breeding can take place in tanks. In addition, the mature shark (around 1-1.5 m in length) responds well for anaesthetics and can take 10 – 15 ml blood for research without significant harm to the fish health.

IgNAR consisted of two parts namely, variable region and constant region of which constant region has five domains. As evolutionary first jawed species to express immunoglobulins (Dooley & Flajnik, 2006), this study reported the presence of all conserved cysteine residues in C region which is also in parallel to human IgG constant region. Data derived from the NGS analysis revealed the presence of two isoforms of IgNAR in brownbanded bamboo shark and around 13 different clusters (chapter 2).

According to the classification of IgNAR based on the position of non-canonical cysteine residues, IgNAR type I, II and IV (IIb) were reported while type III only found in neonates. This classification is based on the presence of extra cysteine residue in different domains of vNAR. Out of this category, type I which consisted of non-canonical cysteine in both FR2 and FR4 regions was only recorded in nurse shark to date (Nuttall et al. 2001; Smith et al. 2012a; Crouch et al. 2013; Barelle & Porter 2015) Among the large set of vNAR amplicon sequences, we have identified one sequence similar to type I IgNAR which is a significant finding. This study also reported novel patterns of IgNAR based on the position of non-canonical cysteine residue in variable domains. Another interesting finding was the presence of cysteine residue in leader sequence which needs to be clarified whether it plays a role in formation of disulfide bond or not. Identification of a variety of sequence types became possible due to construction of vNAR amplicon sequence library. Based on the results from present study, the diversity of vNAR indicates that there is a possibility of changing structural conformations in brownbanded bamboo shark when exposed to antigen. It is also suggestive of the individual variations of responding to antigen. The reduction of the number of unique CDR1 and CDR3 domains along the increase of antigen exposure

might indicate the positive selection of anti-HEL clone during somatic hypermutation (chapter 3).

Affinity maturation in fish (both teleost and cartilaginous species) is believed to be poor due to lack of true germinal centers (Magor 2015). But certain studies on shark immunoglobulins found the significant increase of affinity due to continuous exposure to antigens. However studies on IgNAR showed controversial results indicating species variations such as small spotted catshark did not produce antigen specific IgNAR (Crouch et al. 2013) while nurse shark was a model species for production of antigen specific IgNAR (Dooley et al. 2006) and artificially randomized NAR was synthesized from banded hound shark (Ohtani et al. 2013). This study on brownbanded bamboo shark added another species to the list that have the ability to develop antigen specific IgNAR response and the results widens the current understanding on IgNAR affinity maturation. Our results indicated that minor changes occur in the primary repertoire after exposure of antigens in B lymphocytes which the mutations are scattered in CDR1 and mainly in CDR3 regions (chapter 3).

Most of the studies on shark IgNAR performed sequence analysis and protein analysis separately and according to the methods adopted the relationship between transcriptome and ELISA were unable to determine. In this study vNAR sequence amplification of consecutively immunized shark mRNA revealed the possibility of using mutation frequency to determine antibody expression in addition to affinity maturation. Conventional cloning methods takes time and labor to obtain several unique colonies (Feng et al., 2019). The screening of unique clones by sequencing also takes additional time and labour. This study proposes the NGS based amplicon sequencing

method and molecular screening method to determine highest affinity and specific vNAR gene sequence from a very large sequence pool (chapter 3 & 5).

Transcriptome analysis provides insights to the cellular responses and gene expression profiles. There are few tissues of shark species studied using transcriptome analysis including brownbanded bamboo shark (Hara et al., 2018). PBLs involved in acquired immunity plays a major role in defense system against pathogens. Analysis of immune response of PBLs in bamboo shark have not been studied before and the findings from this study reveals the up-regulation of genes involved in acquired immunity. Transcriptome analysis of Leydig organ and splenic cells also performed and revealed the LO plays major role on aerobic respiration and ATP synthesis (chapter 4).

The ultimate goal of this research was to develop a novel platform for the screening of highly specific vNAR to synthesize monoclonal antibody using CHO cells. The outcomes generated here laid the pathway for future research directing target-based vNAR library generation, screening and development of therapeutic and diagnostic nanobodies (chapter 5).

In conclusion, the results incur from this study suggested that brownbanded bamboo shark is a suitable species for immunization and their diverse vNAR is a possible candidate for the development of immunotherapeutics. The method introduced here can be used for future experiments on immunization and expression using CHO cells.

6.2 Future prospective

Brownbanded bamboo shark IgNAR transcriptome analysis showed that there were several variations at the position of non-canonical cysteine residues. That indicates the structural conformation changes among different types due to formation of disulfide bridges. The structural analysis of candidate vNAR gene against HEL showed the indentation of CDR3 region towards the cryptic region of HEL antigen. The highest number of sequences formed contacts with HEL using cysteine residue and it plays an important role in antigen binding as well as maintaining the structural stability. Hence another immunization trial using bovine serum albumin (BSA) was carried out following the protocols introduced in chapters 3 and 5. Potential anti-BSA vNAR library was created and screened using the spiny dog fish anti-HSA vNAR as template (figure 6.1). This study resulted in formation of strong Ag-Ab complex with Trp, Asp, Tyr and Pro as the prominent amino acids in CDR3.

The candidate vNAR gene sequence derived from immunized bamboo shark was subjected to determine its antigen binding affinity using UCSF Chimera softwre. It revealed the formation of 33 bonds while its anti-BSA nanobody (5VNV) formed only nine contacts at 0.0 A° distance (figure 6.2). Therefore, the candidate gene has potential binding ability higher than the already established reference sequence indicating the promising application of the novel platform proposed in this study.

This future research will be able to create vNAR nanobody which may bind to human serum albumin to increase serum half-life by reducing glomerular filtration rate due to large size.

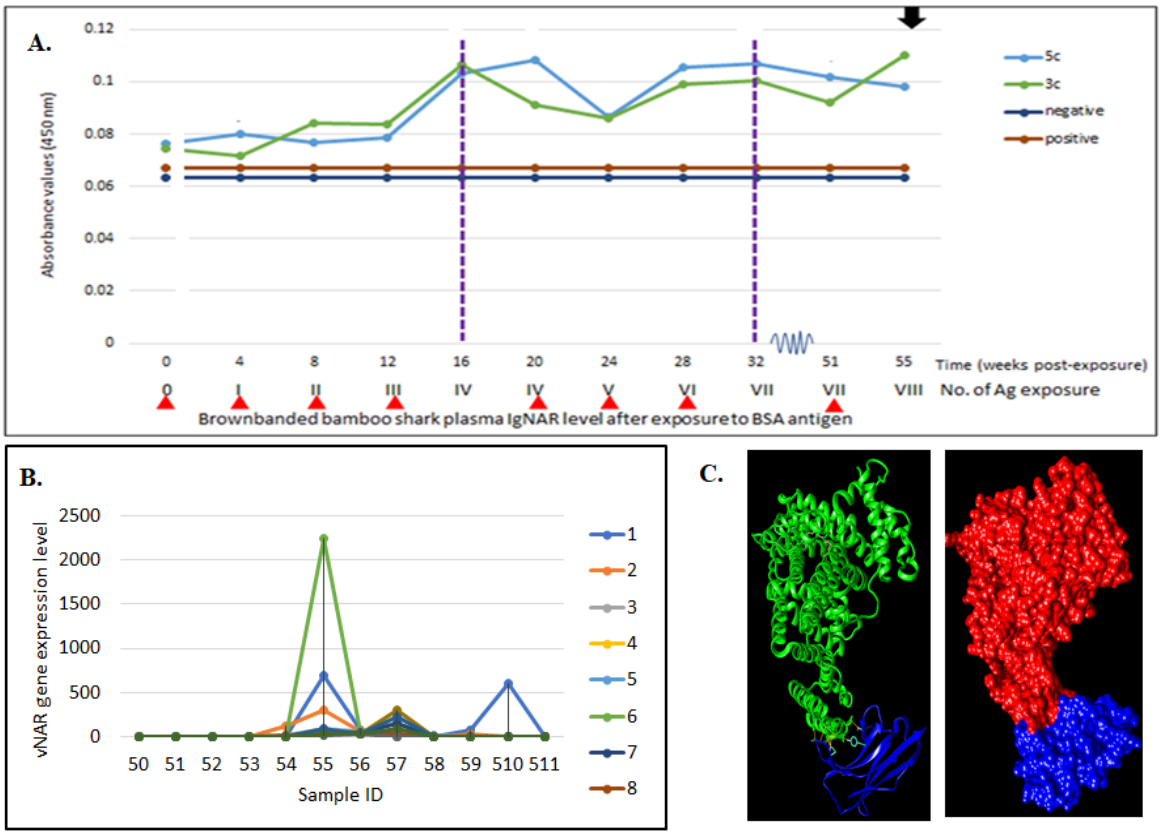


Figure 6.1. Key results generated by the ongoing study on developing anti-BSA vNAR nanobody. **A.** ELISA for detection of IgNAR affinity against BSA in two immunized sharks. 5c and 3c are two individual sharks and the red arrow heads indicate the immunizations. Phosphate buffered saline was used as negative control and goat IgG was the positive control. **B**. Eight sequences from shark-5c vNAR library showing the expression values. 50 is the unimmunized sample and consecutive numbering indicates several sample collection points. **C**. Anti-BSA vNAR from brownbanded bamboo shark forming Ag-Ab complex with BSA. Anti-BSA vNAR is in blue colour.

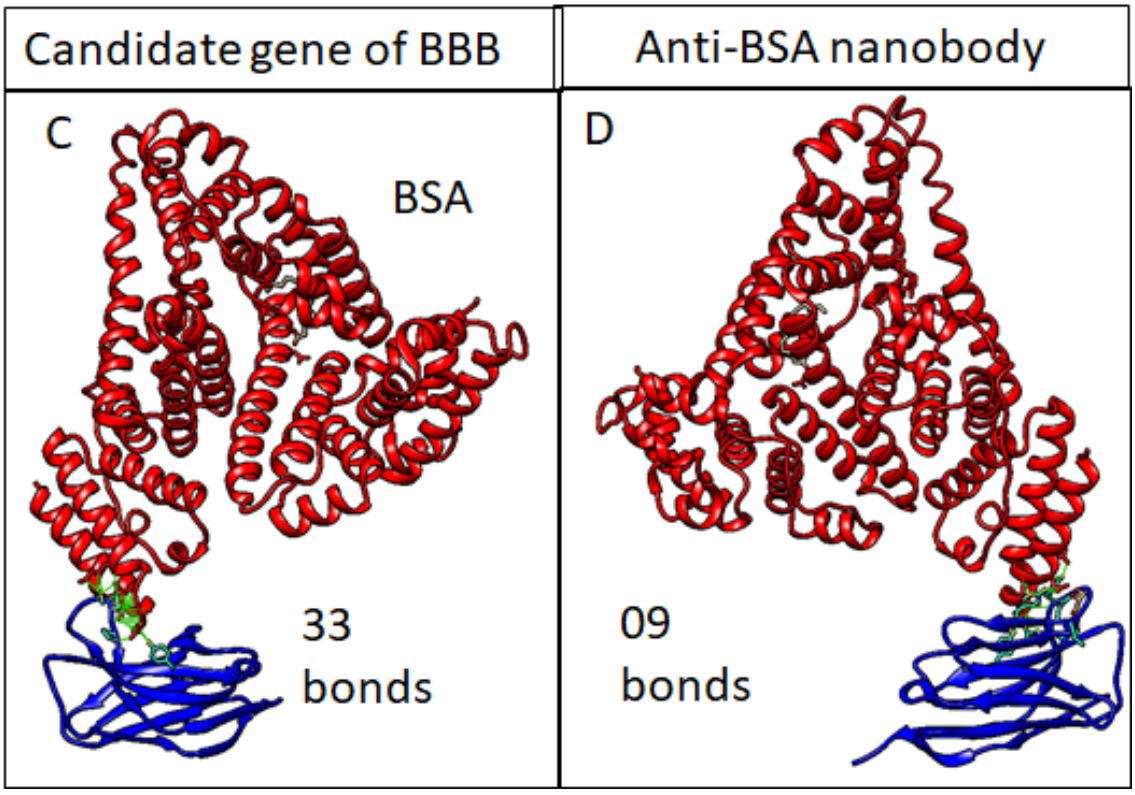
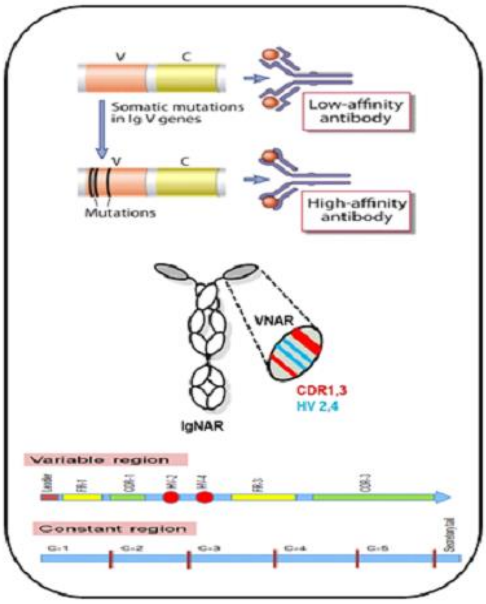
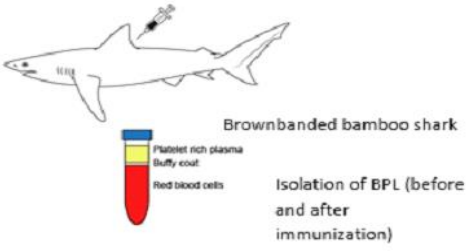


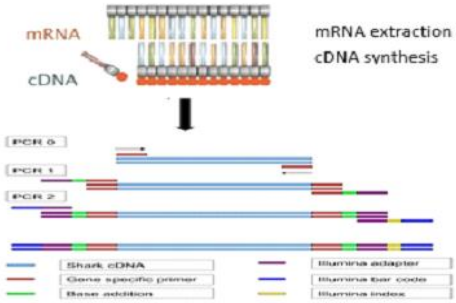
Figure 6.2. Comparison of the binding ability of brownbanded bamboo shark anti-BSA vNAR with anti-BSA nanobody (5VNV).

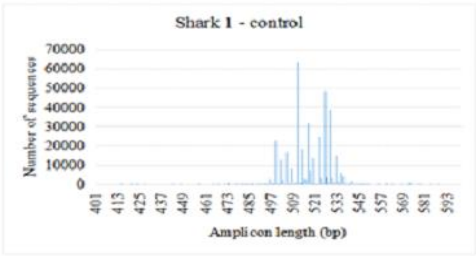
In this study, possibility of affinity maturation and *in vivo* synthesis of antigen specific vNAR was identified. Therefore, similar protocols can be followed to develop further IgNAR antibodies against specific antigens such as proteins expressed in cancer cells or virulent proteins in pathogenic microorganisms. Synthesis of brownbanded bamboo shark origin vNAR monoclonal antibody against such antigens will be useful to reduce adverse effects and side reactions due to higher specificity and ability to reach tiny cellular spaces. Finally, data presented in this study will be useful for the future advancement on immunity related research and developments using sharks.

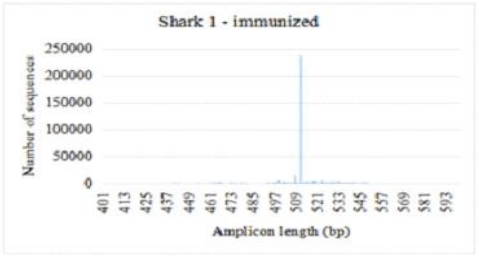
Illustrative summary of the research:

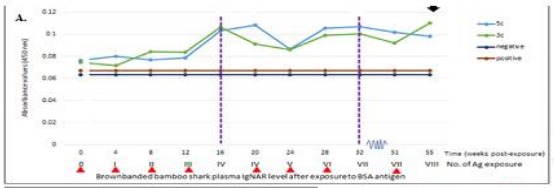


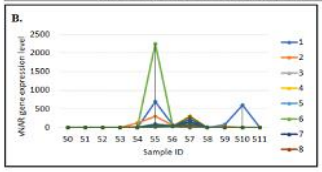


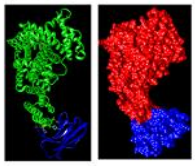


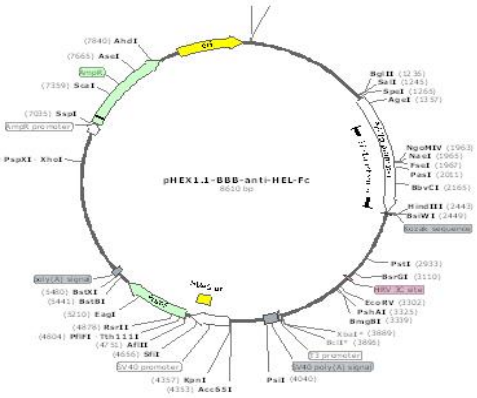


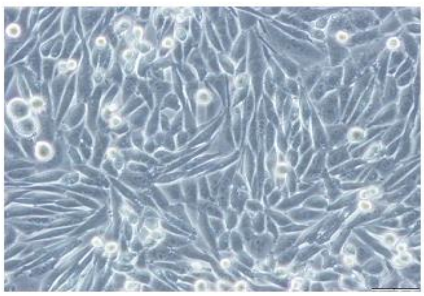


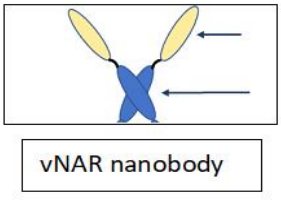












References

Benichou, J., Ben-Hamo, R., Louzoun., Y. & Efroni, S. 2012. Rep-Seq: Uncovering the immunological repertoire through next-generation sequencing. *Immunology*, 135(3), 183–191.

Bolger, A.M., Lohse, M. & Usadel, B. 2014. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics*, 30(15), 2114–2120.

Bray N.L., Pimentel H., Melsted P. and Pachter L., 2016 Near-optimal probabilistic RNA seq quantification. *Nat. Biotechnol*., 34 (5), 525-527.

Crouch, K., Smith, L.E., Williams, R., Cao, W., Lee, M., Jensen, A. & Dooley, H. 2013. Humoral immune response of the small-spotted catshark, Scyliorhinus canicula. *Fish and Shellfish Immunology*, 34(5), 1158–1169.

De los Rios, M., Criscitiello, M. & Smider, V. 2015. Structural and genetic diversity in antibody repertoires from diverse species. *Current Opinion in Structural Biology*, 33, 27-41.

Diaz, M., Greenberg, A.S. & Flajnik, M.F., 1998. Somatic hypermutation of the new antigen receptor gene (NAR) in the nurse shark does not generate the repertoire: possible role in antigen-driven reactions in the absence of germinal centers. *Proceedings of the National Academy of Sciences of the United States of America*, 95(24), 14343–14348.

Diaz, M., Stanfield, R.L., Greenberg, A.S. & Flajnik, M.F., 2002. Structural analysis, selection, and ontogeny of the shark new antigen receptor (IgNAR): Identification of a new locus preferentially expressed in early development. *Immunogenetics*,

54(7), 501–512.

Dooley, H., Flajnik, M.F. & Porter, A.J., 2003. Selection and characterization of naturally occurring single-domain (IgNAR) antibody fragments from immunized sharks by phage display. *Molecular Immunology*, 40(1), 25–33.

Dooley, H. & Flajnik, M.F., 2005. Shark immunity bites back: Affinity maturation and memory response in the nurse shark, Ginglymostoma cirratum. *European Journal of Immunology*, 35(3), 936–945.

Dooley, H. & Flajnik, M.F., 2006. Antibody repertoire development in cartilaginous fish. *Developmental and Comparative Immunology*, 30(1–2), 43–56.

Dooley, H., Stanfield, R.L., Brady R.A. & Flajnik M.F., 2006. First molecular and biochemical analysis of in vivo affinity maturation in an ectothermic vertebrate. *Proceedings of the National Academy of Sciences of the United States of America*, 103(6), 1846–51.

Feige, M.J., Grawert, M.A., Marcinowski, M., Henning, J., Behnke, J., Auslander, D.Herold, E.M., Peschek, J., Castro, C.D., Flajnik, M., Hendershot, L.M., Sattler, M., Groll, M. & Buchner, J., 2014. The structural analysis of shark IgNAR antibodies reveals evolutionary principles of immunoglobulins. *Proceedings of the National Academy of Sciences of the United States of America*, 111(22), 8155–60.

Felsenstein, J., 1985. Confidence limits on phylogenies: An approach using the bootstrap. *Evolution*, 39, 783-791

Feng, M., Bian, H., Wu, X., Fu, T., Fu, Y., Hong, J., Flemming, B.D., Flajnik, M.F. & Ho, M., 2019. Construction and next-generation sequencing analysis of a large phage-displayed VNAR single-domain antibody library from six naïve nurse sharks.

Antibody Therapeutics. 2 (1), 1-11.

Flajnik,M.F. Rumfelt, L.L., 2000. The immune system of cartilagenous fish. *Curr. Top. Microbiol. Immunol,* 248, 249-270

Flajnik, M.F., 2002. Comparative analyses of immunoglobulin genes: surprises and portents. *Nature reviews. Immunology*, 2(9), 688–98.

Flajnik, M.F. & Dooley, H., 2009. The generation and selection of single-domain, v region libraries from nurse sharks. *Methods in molecular biology (Clifton, N.J.)*, 562, 71–82.

Flajnik, M.F. & Kasahara, M., 2010. Origin and evolution of the adaptive immune system: genetic events and selective pressures. *Nature reviews. Genetics*, 11(1), 47–59.

Flajnik, M.F., Deschacht, N. & Muyldermans, S., 2011. A case of convergence: Why did a simple alternative to canonical antibodies arise in Sharks and Camels? *PLoS Biology*, 9(8), 1-5.

Gilbert, P.W. & Wood, F.G., 1957. Method of anesthetizing large sharks and rays safely and rapidly. *Science*, 126(3266), 212–213.

Grabherr MG, Haas BJ, Yassour M, Levin JZ, Thompson DA, Amit I, Adiconis X, Fan L, Raychowdhury R, Zeng Q, Chen Z, Mauceli E, Hacohen N, Gnirke A, Rhind N, di Palma F, Birren BW, Nusbaum C, Lindblad-Toh K, Friedman N, Regev A., 2011. Full-length transcriptome assembly from RNA-seq data without a reference genome. *Nat Biotechnol*, 29(7), 644-652.

Greenberg, A.S., Avila, D., Hughes, M., Hughes, A., McKinney, E.C. & Flajnik, M.F., 1995. A new antigen receptor gene family that undergoes rearrangement and

extensive somatic diversification in sharks. *Nature*, 374(6518), 168–73.

Griffiths K, Dolezal O, Parisi K, Angerosa J, Dogovski C, Barraclough M, Sanalla A, Casey J, González I, Perugini M, Nuttall S, Foley M., 2013. Shark variable new antigen receptor (VNAR) single domain antibody fragments: stability and diagnostic applications. *Antibodies*, 2, 66–81.

Hall, T.A., 1999. BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. *Nucl. Acids Symp*, 41, .95–98.

Hamers-Casterman, C., Atarhouch, T., Muyldemans, S., Robinson, G., Hamers, C., Songa, E.B., Bendahman N. and Hamers, R., 1993. Naturally occurring antibodies devoid of light chains. *Nature*, 363(6428), 446–448.

Hara, Y., Yamaguchi, K., Onimaru, K. *et al.* 2018. Shark genomes provide insights into elasmobranch evolution and the origin of vertebrates. *Nat Ecol Evol* 2**,** 1761–1771

Hinds, K.R. & Litman, G.W., 1986. Major reorganization of immunoglobulin VH segmental elements during vertebrate evolution. *Nature*, 320(6062), 546–9.

Honda, Y., Kondo, H., Caipang, C.M.A., Hirono, I. & Aoki, T., 2010. cDNA cloning of the immunoglobulin heavy chain genes in banded houndshark Triakis scyllium. *Fish and Shellfish Immunology*, 29(5), 854–861.

Honma H., Okabe K. and Chiba A., 1984. Comparative histology of the Leydig and Epigonal orggans in some elasmobranchs. *Japanese Journal of Ichthhyology*, 31 (1), 47-54

Huang, Jinghe; Doria-Rose, Nicole A.; Longo, Nancy S.; Laub, Leo; Lin, Chien-Li; Turk, Ellen; Kang, Byong H.; Migueles, Stephen A.; Bailer, Robert T., 2013. Isolation of human monoclonal antibodies from peripheral blood B cells. *Nature*

Protocols. 8 (10), 1907–1915.

Hinds KR, Litman GW., 1986. Major reorganization of immunoglobulin VH segmental elements during vertebrate evolution. *Nature*. 320, 546–549.

Khan K.H., 2013. Gene expression in mammalian cells and its applications. *Advanced Pharmaceutical Bulletin*. 3(2), 257-263.

Kokubu F, Litman R, Shamblott MJ, Hinds K, Litman GW., 1988. Diverse organization of immunoglobulin VH gene loci in a primitive vertebrate. *EMBO J*., 7, 3413–3422.

Kopsidas, G., Roberts, A.S., Coia, G., Streltsov, V.A. & Nuttall, S.D., 2006. In vitro improvement of a shark IgNAR antibody by Qß replicase mutation and ribosome display mimics in vivo affinity maturation. *Immunology Letters*, 107(2), 163–168.

Kovaleva, M., Ferguson, L., Steven, J., Porter, A. & Barelle, C., 2014. Shark variable new antigen receptor biologics - a novel technology platform for therapeutic drug development. *Expert opinion on biological therapy*, 14(10), 1527–1539.

Kroeker, A.L.; Berard, A.R.; Coombs, K.M., 2012. Fishing beyond the peer: Future omic analyses of virus-host Interactions. *J. Virol. Antivir. Res*. 1(1),.

Kubota, T., Niwa, R., Satoh M., Akinaga S., Shitara, K. & Hanai, N., 2009. Engineered therapeutic antibodies with improved effector functions. *Cancer Science*, 100(9), 1566–1572.

Kumar S., Stecher, G., Tamura, K., 2016. MEGA7: Molecular Evolutionary Genetics Analysis version 7.0 for bigger datasets. *Molecular Biology and Evolution*, 33, 1870-1874

Lee J.S., Park H.J. Kim Y.H. and Lee G.M., 2010. Protein reference mapping of dihydrofolate reductase deficient CHO DG44 cell lines using 2-dimensional electrophoresis. *Proteomics*. 10 (12), 2292-2302.

Leow C.H., Fischer K., Leow C.Y., Cheng Q., Chuah C. and McCarthy J., 2017. Single domain antibodies as new biomarker detectors. *Diagnostics.* 7, 52.

Li, S., Lefrank, M.P., Miles, J.J., Alamyar, E., Giudicelli, V., Durox, P., Freeman, J.D., Corbin, V.D., Scheerlink, J.P., Frohman, M.A., Cameron, P.U., Plebanski, M., Loveland, B., Burrows, S.R., Papenfuss A.T. & Gowans, E.J., 2013. IMGT/HighV QUEST paradigm for T cell receptor IMGT clonotype diversity and next generation repertoire immunoprofiling. *Nat. Commun*. 4:2333

Litman, G.W., Anderson, M.K. & Rast, J.P., 1999. Evolution of antigen binding receptors. *Annual Review of Immunology*, 17(1), 109–147.

Love M.I., Huber W. and Anders S., 2014. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol*. 15, 550.

Magor, B.G., 2015. Antibody Affinity Maturation in Fishes-Our Current Understanding. *Biology*, 4(3), 512–24.

Marra N.J., Richards V.P., Early A., Bogdanowicz S.M., Bitar P.D.P., Stanhope M.J. and Shivji M.S., 2017. Comparative transcriptomics of elasmobranch and teleosts highlight important processes in adaptive immunity and regional endothermy. *BMC Genomics*, 18, 87

Marchalonis, J.J., Schluter, S.F., Bernstein, R.M., Shen S., &. Edmundson, A.B., 1998. Phylogenetic emergence and molecular evolution of the immunoglobulin family, *Adv. Immunol*, 70, 417-506

Matz, H. & Dooley H., 2019. Shark IgNAR-derived binding domains aas potential diagnostic and therapeutic agents. *Developmental and Comparative Immunology*, 90, 100-107.

McCullough, K.C. & Summerfield, A., 2005. Basic Concepts of Immune Response and Defense Development. *ILAR Journal*, 46(3), 230–240.

Miracle, A.L., Anderson, M.K., Litman, R.T., Walsh, C.J., Luer, C.A., Rothenberg, E.V. & Litman, G.W., 2001. Complex expression patterns of lymphocyte-specific genes during the development of cartilaginous fish implicate unique lymphoid tissues in generating an immune repertoire. *International immunology*, 13(4), 567–580.

Morozova, O.; Hirst, M.; Marra, M.A., 2009. Applications of new sequencing technologies for transcriptome analysis. *Annu. Rev. Genom. Hum. Genet*. 10, 135–151.

Müller, M.R., Saunders, K., Grace, C. et. Al., 2012. Improving the pharmacokinetic properties of biologics by fusion to an anti-HSA shark VNAR domain. *mAbs*, 4(6), 673–685.

Mulley J.F., Hargreaves A.D., Hegarty M.J., Heller R.S., Swain M.T., 2014 Transcriptomic analysis of the lesser spotted catshark (Scyliorhinus canicula) pancreas, liver and brain reveals molecular level conservation of vertebrate pancreas function. *BMC Genomics*, 15, 1074.

Mutz, K.O.; Heilkenbrinker, A.; Lönne, M.; Walter, J.G.; Stahl, F. 2013. Transcriptome analysis using next-generation sequencing. *Curr. Opin. Biotechnol*., 24, 22–30.

Nikolova, V., Koo, C., Abrahim, S.A. et al., 2009. Differential roles for membrane-bound and soluble syndecan-1 (CD138) in breast cancer progression. *Carcinogenesis*,

30(3), 397–407.

Nuttall, S.D., Krishnan, U.V., Hattarki, M., De Gori, R., Irving, R.A. & Hudson, P.J., 2001. Isolation of the new antigen receptor from wobbegong sharks, and use as a scaffold for the display of protein loop libraries. *Molecular Immunology*, 38(4), 313–326.

Ogino S., Gulley M.L., Dunnen J.T., Wilson R.B. and the Association for Molecular Pathology Training and Education Committee. 2007. Standard mutation nomenclature in molecular diagnostics. Practical and educational challenges. *Journal of Molecular Diagnostics*, 9(1), 1-7.

Ohtani, M., Hikima, J., Jung, T.S., Kondo, H., Hirono, I. & Aoki, T., 2013. Construction of an Artificially Randomized IgNAR Phage Display Library: Screening of Variable Regions that Bind to Hen Egg White Lysozyme. *Marine Biotechnology*, 15(1), 56–62.

Omasa T, Onitsuka M, Kim W.D., 2010. Cell engineering and cultivation of chinese hamster ovary (CHO) cells. *Curr Pharm Biotechnol*, 11, 233–240.

Oulion S, Debiais-Thibaud M, D'Aubenton-Carafa Y, Thermes C, Da Silva C, Bernard-Samain S, Gavory F, Wincker P, Mazan S, Casane D., 2010. Evolution of Hox gene clusters in gnathostomes: insights from a survey of a shark (*Scyliorhinus canicula*) transcriptome. *Mol Biol Evol*. 27, 2829–2838.

Pettersen E.F., Goddard T.D., Huang C.C., Couch G.S., Greenblatt D.M., Meng E.C., Ferrin T.E., 2004. UCSF Chimera – a visualization system for exploratory research and analysis. *J Comput Chem*, 13, 1605-1612.

Pimentel H., Bray N.L., Puente S., Melsted P. and Pachter L., 2017. Differential analysis

of RNA-seq incorporating quantification uncertainty. *Nat. Methods*. 14 (7), 687-690.

Presta, L.G., 2008. Molecular engineering and design of therapeutic antibodies. *Curr Opin Immunol*, 20(4), 460–470.

Puck T.T., Cieciura S.J. & Robinson A., 1958. Genetics of somatic mammalian cells. III. Long-term cultivation of euploid cells from human and animal subjects. *J Exp Med*. 108(6), 945-956.

Reichert J.M., 2012. Marketed therapeutic antibodies compendium. *MAbs*. 4, 413-415.

Richards V.P., Suzuki H, Stanhope M.J, Shivji M.S. 2013. Characterization of the heart transcriptome of the white shark (*Carcharodon carcharias*). *BMC Genomics*, 14, 697.

Rumfelt, L.L., Avila D., Marilyn D., Barti S.,McKinney E.C. & Flajnik M.A. 2001. A shark antibody heavy chain encoded by a nonsomatically rearranged VDJ is preferentially expressed in early development and is convergent with mammalian IgG. *Proceedings of the National Academy of Sciences of the United States of America*, 98(4), 1775–1780.

Saitou N. and M. Nei, 1987. The neighbor-joining method: A new method for reconstructing phylogenetic trees. *Molecular Biology and Evolution*. 4, 406-425.

Schroeder, H.W. & Cavacini, L., 2010. Structure and Function of Immunoglobulins. *Journal of Allergy and Clinical Immunology*, 125(202) S41-S52.

Seeber, S., Ros, F., Thorey, I., Tiefenthaler, G., Kaluza, K., Lifke, V., Fischer, J., André A., Klostermann, S. & Endl, J. 2014. A Robust High Throughput Platform to Generate Functional Recombinant Monoclonal Antibodies Using Rabbit B Cells

from Peripheral Blood. *PLoS ONE*. 9 (2): e86184.

Simmons D.P., Abregu F.A., Krishnan U.V., Proll D.F., Streltsov V.A., Doughty L, Hattarki M.K. & Nuttall S.D. 2006. Dimerisation strategies for shark IgNAR single domain antibody fragments. *J. of Immunol Methods*. 315(1-2), 171-84.

Skerra A. 2007. Alternative non-antibody scaffolds for molecular recognition. *Curr Opin Biotechnol*, 18, 295-304.

Smith, M., Warmolts, D., Thoney, D. & Hueter, R. 2004. Elasmobranch husbandry manual: Captive care of sharks, rays, and their relatives. *The Elasmobranch Husbandry Manual : Captive Care of Sharks , Rays and their Relatives*, 307-323.

Smith, L.E., Crouch, K., Cao, W. et al., 2012. Characterization of the immunoglobulin repertoire of the spiny dogfish (Squalus acanthias). *Developmental & Comparative Immunology*, 36(4), 665–679.

Stanfield, R.L., Dooley,H., Verdino, P., Flajnik, M.F. & Wilson, I.A., 2007. Maturation of Shark Single-domain (IgNAR) Antibodies: Evidence for Induced-fit Binding. *Journal of Molecular Biology*, 367(2), 358–372.

Streltsov, V.A., Carmichael, J.A. & Nuttall, S.D., 2005. Structure of a shark IgNAR antibody variable domain and modeling of an early-developmental isotype. *Protein science : a publication of the Protein Society*, 14(11), 2901–9.

Sudhakar A., Kumar G & El-Matbouli M. 2018. Transcriptome analysis based on RNA-Seq in understanding pathogenic mechanisms of diseases and immune system of fish: A comprehensive review. *International Journal of Molecular Sciences*. 19, 245

Takagi Y., Kikuchi T., Wada R. and Omasa T. 2017. The enhancement of antibody

concentration and achievement of high cell density CHO cell cultivation by adding nucleoside. *Cytotechnology*. 69, 511-521.

Takechi M, Takeuchi M, Ota K.G, Nishimura O, Mochii M, Itomi K, Adachi N, Takahashi M, Fujimoto S, Tarui H, Okabe M, Aizawa S, Kuratani S. 2011. Overview of the transcriptome profiles identified in hagfish, shark, and bichir: current issues arising from some nonmodel vertebrate taxa. *J Exp Zool Part B Mol Dev Evol*, 316B, 526–546.

Tonegawa, S., 1983. Somatic generation of antibody diversity. *Nature*, 302(5909), 575–581.

Venkatesh B, Lee AP, Ravi V, Maurya AK, Lian MM, Swann JB, Ohta Y, Flajnik MF, Sutoh Y, Kasahara M, Hoon S, Gangu V, Roy SW, Irimia M, Korzh V, Kondrychyn I, Lim ZW, Tay B-H, Tohari S, Kong KW, Ho S, Lorente-Galdos B, Quilez J, Marques-Bonet T, Raney BJ, Ingham PW, Tay A, Hillier LW, Minx P, Boehm T, et al. 2014. Elephant shark genome provides unique insights into gnathostome evolution. *Nature*, 505, 174–179.

Walsh G. 2014. Biopharmaceutical benchmarks. *Nat Biotechnol*, 32, 992–1000.

Wang, F., Ekiert, D.C., Ahmad, I., Yu, W., Zhang, Y., Bazirgan, O., Torkamani, A., Raudsepp, T., Mwangi, W., Criscitiello, M.F., Wilson, I.A., Schultz, P.G. & Smider, V.V. 2013. Reshaping antibody diversity. *Cell*, 153(6), 1379-1393.

Wang, Z.; Gerstein, M.; Snyder, M. 2009. RNA-Seq: A revolutionary tool for transcriptomics. *Nat. Rev. Genet*, 10, 57–63.

Wang Q, Arighi C.N., King B.L., Polson S.W., Vincent J., Chen C, Huang H., Kingham B.F., Page S.T., Rendino M.F., Thomas W.K., Udwary D.W., Wu C.H. 2012.

Community annotation and bioinformatics workforce development in concert-Little Skate Genome Annotation Workshops and Jamborees. Database (Oxford). 2012; 2012:bar064.

Wardemann, H., Yurasov, S., Schaefer, A., Young, James W.; Meffre, E. & Nussenzweig, M.C. 2003. Predominant autoantibody production by early human B cell precursors. *Science*. 301 (5638), 1374–1377.

Warrington, R., Watson, W., Kim H.L. & Antonetti R. 2011. An introduction to immunology and immunopathology. *Allergy, Asthma & Clinical Immunology*, 7(Suppl 1):S1, 1-8.

Waterhouse, A., Bertoni, M., Bienert, S., Studer, G., Tauriello, G., Gumienny, R., Heer, F.T., de Beer, T.A.P., Rempfer, C., Bordoli, L., Lepore, R., Schwede, T. 2018. SWISS-MODEL: homology modelling of protein structures and complexes. *Nucleic Acids Res*. 46(W1), W296-W303.

Williams, F., A. & Barclay, A.N., 1988. The immunoglobulin superfamily domains for cell surface recognition. *Annual Review of Immunology*, 6, 381–405.

Wu, J., Mao, X., Cai, T., Luo, J., Wei, L. 2006. KOBAS server: a web-based platform for automated annotation and pathway identification. *Nucleic Acids Res*. 34, W720–W724.

Wurm F.M. & Hacker D., 2011. First CHO genome. *Nat Biotechnol*. 29 (8), 718-720.

Xie, C., Mao, X., Huang, J., Ding, Y., Wu, J., Dong, S., Kong, L., Gao, G., Li, C. & Wei, L. 2011. KOBAS 2.0: a web server for annotation and identification of enriched pathways and diseases. *Nucleic Acids Research*, 39, W316–W322.

Zhang, J., Kobert K., Fluori T. & Stamatakis, A. 2014. PEAR: A fast and accurate Illumina Paired-End reAd mergeR. *Bioinformatics,* 30(5), 614–620.

Zielonka, S., Weber, N., Becker, S. et al., 2014. Shark Attack: High affinity binding proteins derived from shark vNAR domains by stepwise in vitro affinity maturation. *Journal of Biotechnology*, 191, 236–245.

Zielonka, S., Empting, M., Grzeschik, J., et al., 2015. Structural insights and biomedical potential of IgNAR scaffolds from sharks. *mAbs*, 7(1), 15–25.

Zuckerkandl E. & Pauling, L. 1965. Evolutionary divergence and convergence in proteins. Edited in Evolving Genes and Proteins by V. Bryson and H.J. Vogel. Academic Press, New York. 97-166

Websites

www.lifescience.toyobo.co.jp

www.almacgroup.com

www.kln.ac.lk/science/depts/zoology

www.bioinformatics.babraham.ac.uk

www.qiagenbioinformatics.com

http://slideplayer.com/slide/10564113/

Annexure 01

Chapter 2: Appendix 01

```
                                10        20        30        40        50        60
                       ....|....|....|....|....|....|....|....|....|....|....|....|
C. punctatum IgNAR-1 CH  ATGAATGGCATACAGCCTTCTCCACCAGTCATCAGTCTCCACTACTCTGCAACTGAAGAA
                          M  N  G  I  Q  P  S  P  P  V  I  S  L  H  Y  S  A  T  E  E
C. punctatum IgNAR-2 CH  ATGAATGGCATACAGCCTTCTCCACCAGTCATCAGTCTCCACTACTCTGCAACTGACGAA
                          M  N  G  I  Q  P  S  P  P  V  I  S  L  H  Y  S  A  T  D  E
                                70        80        90       100       110       120
                       ....|....|....|....|....|....|....|....|....|....|....|....|
C. punctatum IgNAR-1 CH  CAGAGGGCAAACAAATTTCTGCAGCTGATTTGTCTAATTAGCGGATACTATCCTGAAAGA
                          Q  R  A  N  K  F  L  Q  L  I  C  L  I  S  G  Y  Y  P  E  R
C. punctatum IgNAR-2 CH  CAGAGGGCAAACGGATTTCTGCAGCTGATTTGTCTAATTAGCGGATACTATCCTGAAAGA
                          Q  R  A  N  G  F  L  Q  L  I  C  L  I  S  G  Y  Y  P  E  R
                               130       140       150       160       170       180
                       ....|....|....|....|....|....|....|....|....|....|....|....|
C. punctatum IgNAR-1 CH  ATTGCATTGACCTGGCAGAAGAACGGCAATACCATAAGTTCTGGCTTTACAACTACATCG
                          I  A  L  T  W  Q  K  N  G  N  T  I  S  S  G  F  T  T  T  S
C. punctatum IgNAR-2 CH  ATTGCATTGACCTGGCAGAAGAACGGCAATACCATAAGTTCTGGCTTTACAACTACATCT
                          I  A  L  T  W  Q  K  N  G  N  T  I  S  S  G  F  T  T  T  S
                               190       200       210       220       230       240
                       ....|....|....|....|....|....|....|....|....|....|....|....|
C. punctatum IgNAR-1 CH  CCAACGAAAACAGTGACCAATGATTTTAGCTCTACAAGTTTGCTTAAAGTGCCCCTGAAG
                          P  T  K  T  V  T  N  D  F  S  S  T  S  L  L  K  V  P  L  K
C. punctatum IgNAR-2 CH  TCAACGAAAACAGTGACCAATGATTTTAGCTCTACAAGTTTGCTTAAAGTGCCCCTGCAG
                          S  T  K  T  V  T  N  D  F  S  S  T  S  L  L  K  V  P  L  Q
                               250       260       270       280       290       300
                       ....|....|....|....|....|....|....|....|....|....|....|....|
C. punctatum IgNAR-1 CH  GAATGGAACAGCGGCTCTGTGTACAGTTGTCAAGTTTCTCATTCTGCTACCAACAGTAAC
                          E  W  N  S  G  S  V  Y  S  C  Q  V  S  H  S  A  T  N  S  N
```

C. punctatum IgNAR-2 CH **GAATGGAACAGCGGCTCTGTGTACAGTTGTCAAGTTTCTCATTCTGCTACCAACAGTAAC**

E W N S G S V Y S C Q V S H S A T N S N

310 320 330 340 350 360

....|....|....|....|....|....|....|....|....|....|....|....|

C. punctatum IgNAR-1 CH **CAGAAAAAAGAAATTAGTTCAACATCGGAGATAACTGTATTATTGAGGGATCCATTGGTT**

Q K K E I S S T S E I T V L L R D P L V

C. punctatum IgNAR-2 CH **CAGAGAAAAGAAATTAGTTCAACATCGGAGATAACTGTATTATTGAGGGATCCATTGGTT**

Q R K E I S S T S E I T V L L R D P L V

370 380 390 400 410 420

....|....|....|....|....|....|....|....|....|....|....|....|

C. punctatum IgNAR-1 CH **GAAGAAATCTGGATCAATAAAAGCGCCATTCTGGTTTGCGAAGTGCTTTCTACAGTTTCC**

E E I W I N K S A I L V C E V L S T V S

C. punctatum IgNAR-2 CH **GAAGAAATCTGGATCAATAAAAGCGCCATTCTGGTTTGCGAAGTGCTTTCTACAGTTTCC**

E E I W I N K S A I L V C E V L S T V S

430 440 450 460 470 480

....|....|....|....|....|....|....|....|....|....|....|....|

C. punctatum IgNAR-1 CH **AGTGGGGTAGTTATCTCTTGGATGGTGGATGGAAAGGTGCGGAATGAAGACGTTCAAATT**

S G V V I S W M V D G K V R N E D V Q I

C. punctatum IgNAR-2 CH **AGTGGGGTAGTTATCTCTTGGATGGTGGATGGAAAGGTGCGGAATGAAGACGTTCAAATT**

S G V V I S W M V D G K V R N E D V Q I

490 500 510 520 530 540

....|....|....|....|....|....|....|....|....|....|....|....|

C. punctatum IgNAR-1 CH **GAAGCAGCGAACATAAACCAATATGTGACAATCAGCCGCTTGACTAGCACCGTGGACGAG**

E A A N I N Q Y V T I S R L T S T V D E

C. punctatum IgNAR-2 CH **GAAGCAGCGAACATAAACCAATATGTGACAATCAGCCGCTTGACTAGCACCGTGGACGAG**

E A A N I N Q Y V T I S R L T S T V D E

550 560 570 580 590 600

....|....|....|....|....|....|....|....|....|....|....|....|

C. punctatum IgNAR-1 CH **TGGCAGAGTGGGGTGGAATACACATGTTCCGCAAAACAGGATCAATCATCACCCGCAATA**

W Q S G V E Y T C S A K Q D Q S S P A I

C. punctatum IgNAR-2 CH **TGGCAGAGTGGGGTGGAATACACATGTTCCGCAAAACAGGATAAATCATCACCAGCAATA**

W Q S G V E Y T C S A K Q D K S S P A I

```
                                610       620       630       640       650       660
                         ....|....|....|....|....|....|....|....|....|....|....|....|
C. punctatum IgNAR-1 CH  TCAAAACGAACAAGAATGGCAAGAGTCAAATCAATGAAGCCATATATCCGCCTCCTGCCT
                          S  K  R  T  R  M  A  R  V  K  S  M  K  P  Y  I  R  L  L  P
C. punctatum IgNAR-2 CH  TCAAAACGAACAAGAATGGCAAGAGTCAAATCAATGAAGCCATATATCCGCCTCCTGCCT
                          S  K  R  T  R  M  A  R  V  K  S  M  K  P  Y  I  R  L  L  P
                                670       680       690       700       710       720
                         ....|....|....|....|....|....|....|....|....|....|....|....|
C. punctatum IgNAR-1 CH  CCATCACCAGAAGAAATTAAGAACAGCAGCACTGCTACGCTCACATGTTTGCTAAGAGGG
                          P  S  P  E  E  I  K  N  S  S  T  A  T  L  T  C  L  L  R  G
C. punctatum IgNAR-2 CH  CCATCACCAGAAGAAATTAAGAACAGCAGCACTGCTACGCTCACATGTTTGCTAAGAGGG
                          P  S  P  E  E  I  K  N  S  S  T  A  T  L  T  C  L  L  R  G
                                730       740       750       760       770       780
                         ....|....|....|....|....|....|....|....|....|....|....|....|
C. punctatum IgNAR-1 CH  TTCTATCCTGACAAAATACGCGTTTCCTGGGAGAAAGACGGAGCTTCTGTCAACGTGAAC
                          F  Y  P  D  K  I  R  V  S  W  E  K  D  G  A  S  V  N  V  N
C. punctatum IgNAR-2 CH  TTCTATCCTGACAAAATACGCGTTTCCTGGGAGAAAGACGGAGCTTCTGTCAACGTGAAC
                          F  Y  P  D  K  I  R  V  S  W  E  K  D  G  A  S  V  N  V  N
                                790       800       810       820       830       840
                         ....|....|....|....|....|....|....|....|....|....|....|....|
C. punctatum IgNAR-1 CH  ATCACCAATTTCCCCACTGCTCTGGAACAGGACCTGACCTACAGCACACGGAGCCTCCTC
                          I  T  N  F  P  T  A  L  E  Q  D  L  T  Y  S  T  R  S  L  L
C. punctatum IgNAR-2 CH  ATCACCAATTTCCCCACCGCTCTGGAACAGGACCTGACCTACAGCACACGGAGCCTCCTC
                          I  T  N  F  P  T  A  L  E  Q  D  L  T  Y  S  T  R  S  L  L
                                850       860       870       880       890       900
                         ....|....|....|....|....|....|....|....|....|....|....|....|
C. punctatum IgNAR-1 CH  GTTTTACCTGCAGCGGAATGGAAGAGCGGCGCAAAATACACCTGTACTGCCTCGCATCCA
                          V  L  P  A  A  E  W  K  S  G  A  K  Y  T  C  T  A  S  H  P
C. punctatum IgNAR-2 CH  ATTTTACCTGCAGCGGAATGGAAGAGCGGCGCAAAATACACCTGTACTGCCTCGCATCCA
                          I  L  P  A  A  E  W  K  S  G  A  K  Y  T  C  T  A  S  H  P
```

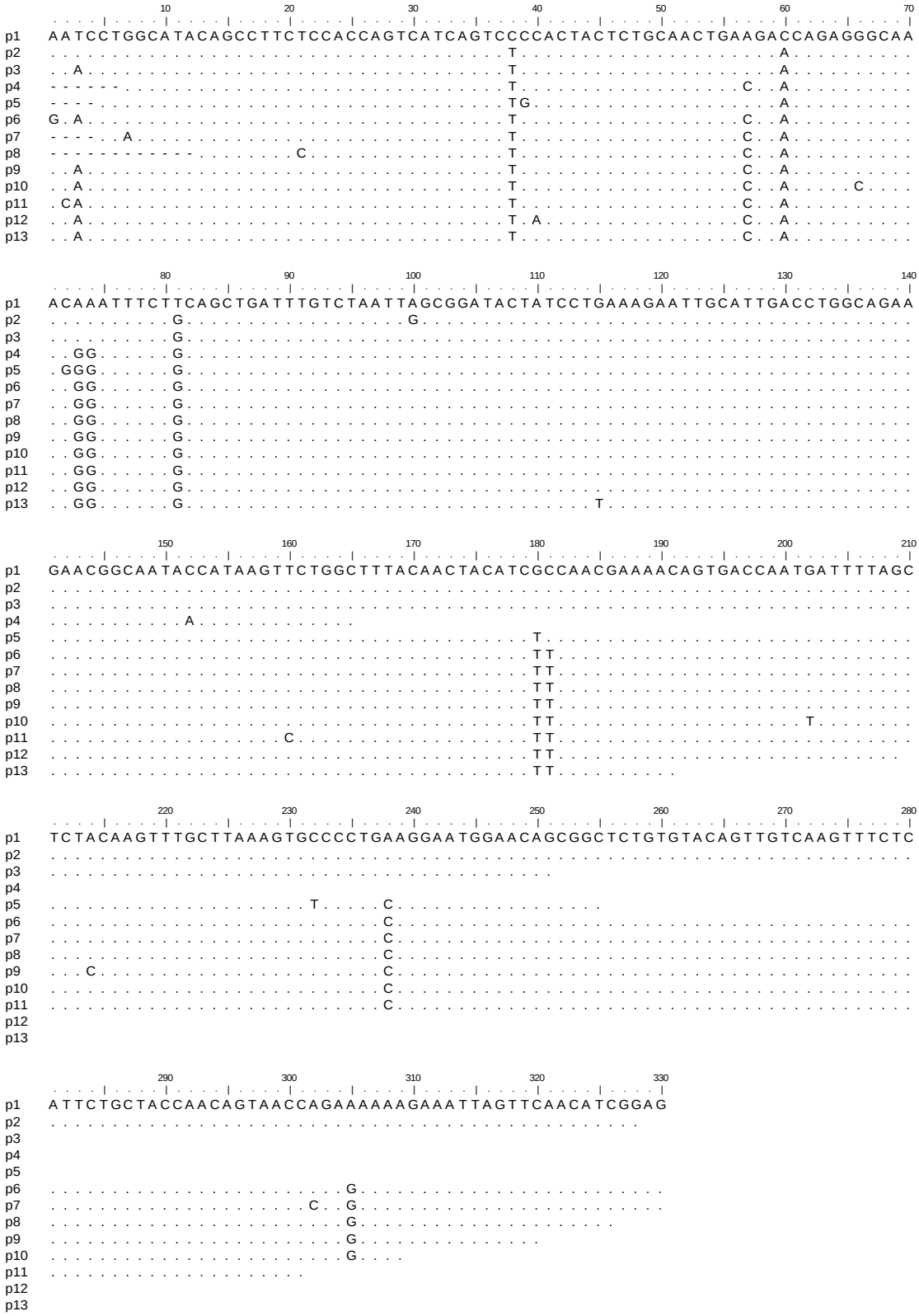
```
                                910       920       930       940       950       960
                           ....|....|....|....|....|....|....|....|....|....|....|....|
C. punctatum IgNAR-1 CH    CCTTCACAGTCCACTGTGAAAAGGGAAATCAGGAGCACGAAAGGTGAATGCCGTCAGACA
                            P  S  Q  S  T  V  K  R  E  I  R  S  T  K  G  E  C  R  Q  T
C. punctatum IgNAR-2 CH    CCTTCACAGTCCACTGTGAAAAGGGAAATCAGGAGCCCGAAAGGTGATTGCCGTCAGACA
                            P  S  Q  S  T  V  K  R  E  I  R  S  P  K  G  D  C  R  Q  T
                                970       980       990       1000      1010      1020
                           ....|....|....|....|....|....|....|....|....|....|....|....|
C. punctatum IgNAR-1 CH    GATATTTCTGTCAATCTACTGAAACCTCCGTTCGAAGAGATTTGGACACAAAAGACAGCG
                            D  I  S  V  N  L  L  K  P  P  F  E  E  I  W  T  Q  K  T  A
C. punctatum IgNAR-2 CH    GATATTTCTGTCAATCTACTGAAACCTCCGTTCGAAGAGATTTGGACACAAAAGACAGCG
                            D  I  S  V  N  L  L  K  P  P  F  E  E  I  W  T  Q  K  T  A
                                1030      1040      1050      1060      1070      1080
                           ....|....|....|....|....|....|....|....|....|....|....|....|
C. punctatum IgNAR-1 CH    ACCATTGTGTGTGAAATCGTTTATAGCAACTTAGAAAACATCAAAGTTTTCTGGCAGGTG
                            T  I  V  C  E  I  V  Y  S  N  L  E  N  I  K  V  F  W  Q  V
C. punctatum IgNAR-2 CH    ACCATTGTGTGTGAAATCGTTTATAGCAACTTAGAAAACATCAAAGTTTTCTGGCAGGTG
                            T  I  V  C  E  I  V  Y  S  N  L  E  N  I  K  V  F  W  Q  V
                                1090      1100      1110      1120      1130      1140
                           ....|....|....|....|....|....|....|....|....|....|....|....|
C. punctatum IgNAR-1 CH    AATGGGGCTGAAAGAACAAAAGGAGTCAAGACAGAAAATCCTGATTGGAACGGAAGTAAA
                            N  G  A  E  R  T  K  G  V  K  T  E  N  P  D  W  N  G  S  K
C. punctatum IgNAR-2 CH    AATGGGGCTGAAAGAACAAAAGGAGTCAAGACAGAAAATCCTGATTGGAACGGAAGTAAA
                            N  G  A  E  R  T  K  G  V  K  T  E  N  P  D  W  N  G  S  K
                                1150      1160      1170      1180      1190      1200
                           ....|....|....|....|....|....|....|....|....|....|....|....|
C. punctatum IgNAR-1 CH    TTCACCATTGTGAGCAAACTTAAAGTCATTGCGTCAGAGTGGGACAGCGGTGCCGAATAT
                            F  T  I  V  S  K  L  K  V  I  A  S  E  W  D  S  G  A  E  Y
C. punctatum IgNAR-2 CH    TTCACCATTGTGAGCAAACTTAAAGTCATTGCGTCAGAGTGGGACAGCGGTGCCGAATAT
                            F  T  I  V  S  K  L  K  V  I  A  S  E  W  D  S  G  A  E  Y
```

```
                                1210      1220      1230      1240      1250      1260
                          ....|....|....|....|....|....|....|....|....|....|....|....|
C. punctatum IgNAR-1 CH   GTTTGCTTGGTGGAAGACAGTGAATTACCAACACCACAGAAAGCTTCTCTCAAGAAAAAA
                           V  C  L  V  E  D  S  E  L  P  T  P  Q  K  A  S  L  K  K  K
C. punctatum IgNAR-2 CH   GTTTGCTTGGTGGAAGACAGTGAATTACCAACACCACAGAAAGCTTCTCTCAAGAAAAAA
                           V  C  L  V  E  D  S  E  L  P  T  P  Q  K  A  S  L  K  K  K
                                1270      1280      1290      1300      1310      1320
                          ....|....|....|....|....|....|....|....|....|....|....|....|
C. punctatum IgNAR-1 CH   ATTGACGACCAAATGCGTCCTCCTCAAGTTTACCTCCTGCAGCCATCGACAGACGAGATT
                           I  D  D  Q  M  R  P  P  Q  V  Y  L  L  Q  P  S  T  D  E  I
C. punctatum IgNAR-2 CH   ATTGACGACCAAATGCGTCCTCCTCAAGTTTACCTCCTGCAGCCATCGACAGACGAGATT
                           I  D  D  Q  M  R  P  P  Q  V  Y  L  L  Q  P  S  T  D  E  I
                                1330      1340      1350      1360      1370      1380
                          ....|....|....|....|....|....|....|....|....|....|....|....|
C. punctatum IgNAR-1 CH   GACGCTGAGAATTCAGCGACTCTGATGTGTCTCGTCACCAGCTTTTACCCAGCTGAGATC
                           D  A  E  N  S  A  T  L  M  C  L  V  T  S  F  Y  P  A  E  I
C. punctatum IgNAR-2 CH   GACGCTGAGAATTCAGCGACTCTGATGTGTCTCGTCACCAGCTTTTACCCAGCTGAGATC
                           D  A  E  N  S  A  T  L  M  C  L  V  T  S  F  Y  P  A  E  I
                                1390      1400      1410      1420      1430      1440
                          ....|....|....|....|....|....|....|....|....|....|....|....|
C. punctatum IgNAR-1 CH   TACATTGGCTGGATGGCCAATGACACCCTTTTGGATACCGGGTACCGAACCCAAGTGGAT
                           Y  I  G  W  M  A  N  D  T  L  L  D  T  G  Y  R  T  Q  V  D
C. punctatum IgNAR-2 CH   TACATTGGCTGGATGGCCAATGACACCCTTTTGGATACCGGGTACCGAACCCAAGTGGAT
                           Y  I  G  W  M  A  N  D  T  L  L  D  T  G  Y  R  T  Q  V  D
                                1450      1460      1470      1480      1490      1500
                          ....|....|....|....|....|....|....|....|....|....|....|....|
C. punctatum IgNAR-1 CH   AGCCAGGCGGGGAATGGTTCCAACTTCGTGACCAGCAGGTTGAGAGTCACGGCGGCGGAA
                           S  Q  A  G  N  G  S  N  F  V  T  S  R  L  R  V  T  A  A  E
C. punctatum IgNAR-2 CH   AGCCAGGCGGGGAATGGTTCCAACTTCGTGACCAGCAGGTTGAGAGTCACGGCGGCGGAA
                           S  Q  A  G  N  G  S  N  F  V  T  S  R  L  R  V  T  A  A  E
```

```
                                1510      1520      1530      1540      1550      1560
                          ....|....|....|....|....|....|....|....|....|....|....|....|
C. punctatum IgNAR-1 CH   TGGAACAATGACACCACTTACTCCTGCTTAGTGGGACACCCATCCCTCAGCCGGGAGTTA
                           W  N  N  D  T  T  Y  S  C  L  V  G  H  P  S  L  S  R  E  L
C. punctatum IgNAR-2 CH   TGGAACAATGACACCACTTACTCCTGCTTAGTGGGACACCCATCCCTCAGCCGGGAGTTA
                           W  N  N  D  T  T  Y  S  C  L  V  G  H  P  S  L  S  R  E  L
                                1570      1580      1590      1600      1610      1620
                          ....|....|....|....|....|....|....|....|....|....|....|....|
C. punctatum IgNAR-1 CH   ATCAGAAGTACAAATAAATCTCATGTAATTGCAGGTATTCCTGAAATAAAAGAGGACCCG
                           I  R  S  T  N  K  S  H  V  I  A  G  I  P  E  I  K  E  D  P
C. punctatum IgNAR-2 CH   ATCAGAAGTACAAATAAATCTCATGTAATTGCAGGTATTCCTGAAATAAAAGAGGACCCG
                           I  R  S  T  N  K  S  H  V  I  A  G  I  P  E  I  K  E  D  P
                                1630      1640      1650      1660      1670      1680
                          ....|....|....|....|....|....|....|....|....|....|....|....|
C. punctatum IgNAR-1 CH   GTAGACTATGACATCGAAGAAAATGAACAAGAACATTTCGATGGAGATGCCAACAGTTCA
                           V  D  Y  D  I  E  E  N  E  Q  E  H  F  D  G  D  A  N  S  S
C. punctatum IgNAR-2 CH   GTAGACTATGACATCGAAGAAAATGAACAAGAACATTTCGATGGAGATGCCAACAGTTCA
                           V  D  Y  D  I  E  E  N  E  Q  E  H  F  D  G  D  A  N  S  S
                                1690      1700      1710      1720      1730      1740
                          ....|....|....|....|....|....|....|....|....|....|....|....|
C. punctatum IgNAR-1 CH   TTATCCATTATGACGTTTGTCGTCCTGTTTATTCTCAGTACGCTGTACAGCACATTTGTA
                           L  S  I  M  T  F  V  V  L  F  I  L  S  T  L  Y  S  T  F  V
C. punctatum IgNAR-2 CH   TTATCCATTATGACGTTTGTCGTCCTGTTTATTCTCAGTACGCTGTACAGCACATTTGTA
                           L  S  I  M  T  F  V  V  L  F  I  L  S  T  L  Y  S  T  F  V
                                1750      1760
                          ....|....|....|....|.
C. punctatum IgNAR-1 CH   ACTGTGCTTAAGGTGAACTGA
                           T  V  L  K  V  N  *
C. punctatum IgNAR-2 CH   ACTGTGCTTAAGGTGAACTGA
                           T  V  L  K  V  N  *
```

Figure A1.1. Nucleotide and amino acid sequences of brownbanded bamboo shark IgNAR type

a)



b)

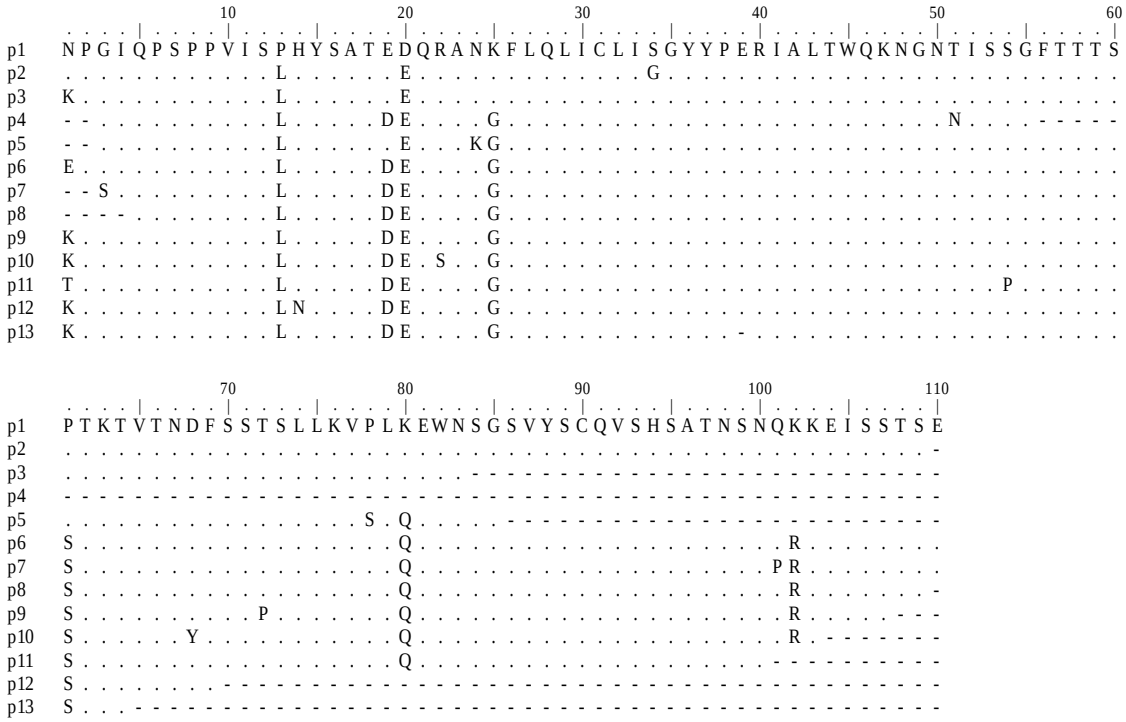


Figure A1.2. Total of 13 unique C1 domain sequences with different combinations of polymorphic residues. **a)** Nucleotide and **b)** amino acid alignment of brownbanded bamboo shark IgNAR. Dots represent identical residues to the first sequence while dash indicates a gap. Single nucleotide polymorphisms were shown at their respective positions.

Table A1.1. Nucleotide variations of brownbanded bamboo shark IgNAR heavy chain constant region obtained from 154 PEAR assembled transcriptome sequences

Mutation No	nucleotide position in IgNAR	A	C	G	T	No. of bases expressed	CH domain	IgNAR CH Sequence	Mutation
1	1	36	0	1	0	37	C1	A	A > G
2	2	34	2	0	1	37	C1	T	T > A, C
3	3	31	0	1	5	37	C1	G	G > A, T
4	4	1	36	0	0	37	C1	A	A > C
5	5	1	39	0	0	40	C1	A	A > C
6	7	1	0	44	0	45	C1	G	G > A
7	21	0	1	0	49	50	C1	T	T > C
8	38	0	1	0	54	55	C1	T	T > C
9	39	0	53	2	0	55	C1	C	C > G
10	40	1	54	0	0	55	C1	C	C > A
11	57	19	40	0	0	59	C1	A	A > C
12	60	58	1	0	0	59	C1	A	A > C
13	66	0	1	55	0	56	C1	G	G > C
14	72	0	54	2	0	56	C1	C	C > G
15	73	17	0	37	0	54	C1	A	A > G
16	74	17	0	37	0	54	C1	A	A > G
17	81	0	0	51	1	52	C1	G	G > T
18	100	49	0	1	0	50	C1	A	A > G
19	115	0	0	47	1	48	C1	G	G > T
20	152	1	43	0	0	44	C1	C	C > A
21	160	0	1	0	43	44	C1	T	T > C

22	180	0	0	15	26	41	C1	G	G > T
23	181	0	17	0	24	41	C1	C	C > T
24	202	0	0	31	1	32	C1	G	G > T
25	214	28	1	0	0	29	C1	A	A > C
26	232	0	27	0	1	28	C1	C	C > T
27	238	5	19	0	0	24	C1	C	C > A
28	302	15	1	0	0	16	C1	A	A > C
29	305	6	0	8	0	14	C1	A	A > G
30	381	9	0	1	0	10	C2	A	A > G
31	383	1	0	8	0	9	C2	G	G > A
32	384	0	8	0	1	9	C2	C	C > T
33	389	0	1	0	8	9	C2	T	T > C
34	399	0	8	0	1	9	C2	C	C > T
35	420	1	7	0	0	8	C2	C	C > A
36	421	6	0	1	0	7	C2	A	A > G
37	422	0	1	6	0	7	C2	G	G > C
38	427	1	0	6	0	7	C2	G	G > A
39	432	0	1	0	6	7	C2	T	T > C
40	453	6	0	1	0	7	C2	A	A > G
41	459	0	1	5	0	6	C2	G	G > C
42	467	5	0	1	0	6	C2	A	A > G
43	470	5	0	1	0	6	C2	A	A > G
44	479	0	1	0	5	6	C2	T	T > C
45	483	5	0	1	0	6	C2	A	A > G
46	489	0	0	5	1	6	C2	G	G > T
47	491	5	0	0	1	6	C2	A	A > T

48	492	0	5	0	1	6	C2	C	C > T
49	493	5	1	0	0	6	C2	A	A > C
50	494	1	0	0	5	6	C2	T	T > A
51	498	0	5	0	1	6	C2	C	C > T
52	502	0	1	0	5	6	C2	T	T > C
53	503	5	0	1	0	6	C2	A	A > G
54	505	0	1	5	0	6	C2	G	G > C
55	510	5	1	0	0	6	C2	A	A > C
56	514	5	0	1	0	6	C2	A	A > G
57	519	0	5	0	1	6	C2	C	C > T
58	530	0	5	1	0	6	C2	C	C > G
59	537	0	5	0	1	6	C2	C	C > T
60	567	0	2	0	5	7	C2	T	T > C
61	570	0	4	0	2	6	C2	C	C > T
62	579	0	2	2	0	4	C2	G	G > C
63	580	2	0	2	0	4	C2	G	G > A
64	582	2	0	0	2	4	C2	T	T > A
65	583	1	3	0	0	4	C2	C	C > A
66	592	2	22	0	0	24	C2	C	C > A
67	594	4	20	0	0	24	C2	C	C > A
68	598	22	0	2	0	24	C2	A	A > G
69	601	0	1	0	23	24	C2	T	T > C
70	603	23	0	1	0	24	C2	A	A > G
71	624	21	0	0	1	22	C2	A	A > T
72	628	21	0	2	0	23	C3	A	A > G
73	631	0	2	0	22	24	C3	T	T > C

74	634	22	0	2	0	24	C3	A	A > G
75	635	2	0	0	22	24	C3	T	T > A
76	643	0	2	0	22	24	C3	T	T > C
77	646	22	2	0	0	24	C3	A	A > C
78	650	0	0	24	1	25	C3	G	G > T
79	670	2	0	25	0	27	C3	G	G > A
80	672	26	2	0	0	28	C3	A	A > C
81	674	28	1	0	0	29	C3	A	A > C
82	675	27	0	2	0	29	C3	A	A > G
83	676	27	0	2	0	29	C3	A	A > G
84	678	2	0	0	27	29	C3	T	T > A
85	679	27	2	0	0	29	C3	A	A > C
86	681	2	0	27	0	29	C3	G	G > A
87	685	27	0	2	0	29	C3	A	A > G
88	689	0	2	27	0	29	C3	G	G > C
89	691	27	0	0	2	29	C3	A	A > T
90	699	1	0	25	2	28	C3	G	G > T, A
91	708	1	0	0	27	28	C3	T	T > A
92	711	0	0	29	1	30	C3	G	G > T
93	712	3	27	0	0	30	C3	C	C > A
94	720	3	0	27	0	30	C3	G	G > A
95	741	2	27	0	0	29	C3	C	C > A
96	745	2	1	0	26	29	C3	T	T > A, C
97	746	2	24	0	3	29	C3	C	C > T, A
98	773	27	0	2	0	29	C3	A	A > G
99	775	2	0	27	0	29	C3	G	G > A

100	776	2	0	0	27	29	C3	T	T > A
101	783	0	30	0	1	31	C3	C	C > T
102	798	0	11	0	22	33	C3	T	T > C
103	811	1	0	33	0	34	C3	G	G > A
104	812	32	0	1	0	33	C3	A	A > G
105	814	0	31	0	1	32	C3	C	C > T
106	838	1	25	0	0	26	C3	C	C > A
107	841	14	0	12	0	26	C3	G	G > A
108	844	0	0	1	25	26	C3	T	T > G
109	847	1	25	0	0	26	C3	C	C > A
110	854	0	25	0	1	26	C3	C	C > T
111	866	0	0	27	1	28	C3	G	G > T
112	869	0	0	27	1	28	C3	G	G > T
113	870	1	26	0	0	27	C3	C	C > A
114	876	26	0	0	1	27	C3	A	A > T
115	891	1	27	0	0	28	C3	C	C > A
116	893	0	27	0	1	28	C3	C	C > T
117	895	1	27	0	0	28	C3	C	C > A
118	902	6	23	0	0	29	C3	C	C > A
119	909	0	0	27	1	28	C3	G	G > T
120	919	27	1	0	0	28	C3	A	A > C
121	921	25	1	1	0	27	C3	A	A > C, G
122	922	25	2	0	0	27	C3	A	A > C
123	923	0	1	26	0	27	C3	G	G > C
124	924	1	0	26	0	27	C3	G	G > A
125	926	21	0	0	6	27	C3	A	A > T

126	927	21	6	0	0	27	C3	A	A > C
127	930	6	21	0	0	27	C3	C	C > A
128	933	0	1	24	0	25	C3	G	G > C
129	934	24	1	0	0	25	C3	A	A > C
130	935	1	0	24	0	25	C3	G	G > A
131	936	1	23	0	1	25	C3	C	C > A, T
132	937	18	7	0	0	25	C3	A	A > C
133	943	1	0	26	0	27	C3	G	G > A
134	944	1	0	25	0	26	C3	G	G > A
135	948	19	0	0	7	26	C3	A	A > T
136	953	1	0	24	1	26	C4	G	G > A, T
137	958	25	2	0	0	27	C4	A	A > C
138	970	0	1	29	0	30	C4	G	G > C
139	975	0	2	0	28	30	C4	T	T > C
140	976	0	29	0	1	30	C4	C	C > T
141	980	0	0	1	30	31	C4	T	T > G
142	988	0	28	2	0	30	C4	C	C > G
143	993	0	27	0	2	29	C4	C	C > T
144	1020	1	0	28	0	29	C4	G	G > A
145	1026	0	1	0	26	27	C4	T	T > C
146	1032	0	1	0	26	27	C4	T	T > C
147	1033	1	0	26	0	27	C4	G	G > A
148	1043	25	1	0	0	26	C4	A	A > C
149	1044	0	1	0	25	26	C4	T	T > C
150	1050	0	25	1	0	26	C4	C	C > G
151	1055	25	0	1	0	26	C4	A	A > G

152	1065	17	8	0	1	26	C4	A	A > C, T
153	1069	0	0	2	24	26	C4	T	T > G
154	1075	1	24	0	0	25	C4	C	C > A
155	1077	1	0	23	0	24	C4	G	G > A
156	1080	4	0	20	0	24	C4	G	G > A
157	1083	0	1	0	23	24	C4	T	T > C
158	1087	0	0	23	1	24	C4	G	G > T
159	1092	22	0	1	0	23	C4	A	A > G
160	1097	1	22	0	0	23	C4	C	C > A
161	1098	22	1	0	0	23	C4	A	A > C
162	1108	18	0	1	0	19	C4	A	A > G
163	1109	18	0	0	1	19	C4	A	A > T
164	1113	18	0	1	0	19	C4	A	A > G
165	1158	19	1	0	0	20	C4	A	A > C
166	1165	1	0	20	0	21	C4	G	G > A
167	1168	20	0	0	1	21	C4	A	A > T
168	1170	0	0	1	20	21	C4	T	T > G
169	1212	0	1	22	0	23	C4	G	G > C
170	1214	18	0	5	0	23	C4	A	A > G
171	1222	0	1	20	0	21	C4	G	G > C
172	1230	16	0	0	4	20	C4	A	A > T
173	1237	0	13	4	0	17	C4	C	C > G
174	1238	16	0	0	1	17	C4	A	A > T
175	1239	1	0	16	0	17	C4	G	G > A
176	1244	1	13	0	3	17	C4	C	C > T, A
177	1245	0	1	0	15	16	C4	T	T > C

178	1249	1	15	0	0	16	C4	C	C > A
179	1255	15	1	0	0	16	C4	A	A > C
180	1257	13	0	3	0	16	C4	A	A > G
181	1258	14	0	1	0	15	C4	A	A > G
182	1259	14	1	0	0	15	C4	A	A > C
183	1262	0	1	0	14	15	C4	T	T > C
184	1266	0	14	0	1	15	C4	C	C > T
185	1268	13	0	1	0	14	C5	A	A > G
186	1269	0	13	0	1	14	C5	C	C > T
187	1276	0	13	0	1	14	C5	C	C > T
188	1293	0	13	0	1	14	C5	C	C > T
189	1302	0	0	13	1	14	C5	G	G > T
190	1311	12	0	0	1	13	C5	A	A > T
191	1324	1	0	13	0	14	C5	G	G > A
192	1326	0	1	0	13	14	C5	T	T > C
193	1377	0	1	12	0	13	C5	G	G > C
194	1393	13	1	0	0	14	C5	A	A > C
195	1418	0	13	0	1	14	C5	C	C > T
196	1491	3	0	14	0	17	C5	G	G > A
197	1495	3	0	14	0	17	C5	G	G > A
198	1497	0	3	14	0	17	C5	G	G > C
199	1558	0	1	0	18	19	C5	T	T > C

Table A1.2. Total expression of amino acid residues and their percentages of brownbanded bamboo shark secretory form of IgNAR heavy chain constant region

Amino acid	Number of expressions	Percentage of the amino acid residues
Ala A	31	5.28
Cys C	11	1.87
Asp D	26	4.43
Glu E	41	6.98
Phe F	14	2.39
Gly G	23	3.92
His H	6	1.02
Ile I	39	6.64
Lys K	35	5.96
Leu L	46	7.84
Met M	8	1.36
Asn N	32	5.45
Pro P	29	4.94
Gln Q	23	3.92
Arg R	22	3.75
Ser S	68	11.58
Thr T	53	9.03
Val V	47	8.01
Trp W	13	2.21
Tyr Y	19	3.24