



Publicly Accessible Penn Dissertations

---


2016

## The Role Of Antibody Subclass In The Pathogenesis Of Pemphigus Vulgaris

Eric Milan Mukherjee

University of Pennsylvania, [eric.mukherjee@gmail.com](mailto:eric.mukherjee@gmail.com)

Follow this and additional works at: <https://repository.upenn.edu/edissertations>

 Part of the [Allergy and Immunology Commons](#), [Bioinformatics Commons](#), [Genetics Commons](#), [Immunology and Infectious Disease Commons](#), and the [Medical Immunology Commons](#)

---

### Recommended Citation

Mukherjee, Eric Milan, "The Role Of Antibody Subclass In The Pathogenesis Of Pemphigus Vulgaris" (2016). *Publicly Accessible Penn Dissertations*. 2709.  
<https://repository.upenn.edu/edissertations/2709>

This paper is posted at ScholarlyCommons. <https://repository.upenn.edu/edissertations/2709>  
For more information, please contact [repository@pobox.upenn.edu](mailto:repository@pobox.upenn.edu).

---

# The Role Of Antibody Subclass In The Pathogenesis Of Pemphigus Vulgaris

## Abstract

A marvel of evolution, the adaptive immune system has the capacity to respond to almost any foreign antigen in a highly specific manner. Antibodies, Y-shaped glycoproteins containing both diverse variable regions responsible for antigen binding and constant regions responsible for effector function, are a key part of this capacity. However, this vast diversity comes with several drawbacks, one of which is the fact that the immune system can deleteriously respond to self-antigens. The focus of this thesis is to characterize the role of class-switching (the changing of antibody constant regions) in the pathogenesis of autoimmune disease, and in particular to trace the lineage of antigen-specific autoreactive B cells by analyzing clonal relationships between antibodies of different constant regions. Analyzing such lineages has the potential to shed light on mechanisms of autoantibody-mediated disease pathogenesis, leading to better understanding of autoimmunity and better therapeutics.

The work presented in this thesis focuses on pemphigus vulgaris, or PV, a model antibody-mediated autoimmune disease characterized by a response to the cell adhesion protein desmoglein (Dsg) 3, which holds keratinocytes together in the epidermis. An enigmatic feature of this disease is the predominance of antibodies from the IgG4 subclass during active disease, which ordinarily appears to have few effector functions and may serve as a “brake” on the immune system in the setting of continuous stimulation by antigen. PV patients also display autoantibodies of the IgG1 subclass during disease and remission, but the relationship between IgG1 and IgG4 in the disease is unclear. Because the majority of cases of PV also harbor anti-Dsg antibodies of the IgA1 and IgA2 subclasses, we sought to determine the relationships between autoantibodies belonging to each of these subclasses. First, we address whether the same anti-Dsg variable region, grafted onto either IgG1 or IgG4 constant regions, can show differing affinity or pathogenicity, in order to determine whether antibody subclass is directly modulating pathogenic effect (chapter 2). Finding that the subclass has very little effect on antibody affinity, pathogenicity, or epitope preference, we then sought to determine whether B cells expressing autoantibodies of different subclasses share lineages, indicating common pathways of development (chapter 3). Using a combination of antigen-specific antibody cloning through phage display, and next-generation sequencing of subclass specific repertoires in a panel of PV patients, we managed to trace 80 lineages of anti-Dsg B cells across all four subclasses tested. In particular, we found that anti-Dsg IgG4 B cells, which are believed to be central to disease pathogenesis, tended to not share lineages with other subclasses, and in general do not appear to share a precursor-product relationship with anti-Dsg IgG1 B cells. We have also found that anti-Dsg IgA1 and IgA2 were tightly related and often arose directly from IgG precursors. These findings are key to understanding the role of class-switching in the pathogenesis of PV, and may shed light on the class-switch mechanisms driving other autoimmune diseases and states of chronic antigen stimulation.

## Degree Type

Dissertation

## Degree Name

Doctor of Philosophy (PhD)

## Graduate Group

Cell & Molecular Biology

## First Advisor

Aimee S. Payne

---

**Keywords**

Antibody repertoires, Autoimmunity, Dermatology, Pemphigus Vulgaris

**Subject Categories**

Allergy and Immunology | Bioinformatics | Genetics | Immunology and Infectious Disease | Medical Immunology

THE ROLE OF ANTIBODY SUBCLASS IN THE PATHOGENESIS OF PEMPHIGUS VULGARIS

Eric Milan Mukherjee

A DISSERTATION

in

Cell and Molecular Biology

Presented to the Faculties of the University of Pennsylvania

in

Partial Fulfillment of the Requirements for the

Degree of Doctor of Philosophy

2018

Supervisor of Dissertation

---

Aimee S. Payne, MD, PhD  
Albert M. Kligman Associate Professor of Dermatology

Graduate Group Chairperson

---

Daniel Kessler  
Associate Professor of Cell and Developmental Biology

Dissertation Committee

John Seykora, MD, PhD, Associate Professor of Dermatology  
Eline Luning Prak, MD, PhD, Associate Professor of Pathology and Laboratory Medicine  
Donald L. Siegel, MD, PhD, Professor of Pathology and Laboratory Medicine  
David Michael Allman, PhD, Associate Professor of Pathology and Laboratory Medicine

## ACKNOWLEDGEMENTS

“No man is an island.” –John Donne, *Devotions Upon Emergent Occasions*

To my parents, who gave much more than their genes.

To my aunt and uncle (Sephali and Ranjit Das), who always gave their love, support, and excellent Christmas presents.

To my cousin (Arijit Das), who unusually gave me a name easily pronounceable by Americans.

To Mr. Mike Steinert and Mrs. Elaine Pardee, my high school biology teachers, who taught me how to love this subject.

To Dr. Arupa Ganguly, who gave me my first research job as a high school student.

To Skip Brass, Maggie Krall, and Ben Stanger, for their support during easy and hard times.

To John Stanley, for his mentorship, and teaching me that perfect is the enemy of good.

To Aimee Payne, for her guidance, encouragement, kindness, and patience.

To my thesis committee, for their suggestions, critical eye, and encouragement.

To my labmates, teammates, and friends, for their support over the years.

To Mathew Crawford, who taught me how to think creatively.

To the Helfand Lab at Brown, especially Jason Wood, who taught me how to think scientifically.

To Jerry Vinokurov, who taught me how to think enjoyably.

To Saajid Moyeen, Patrick Liao, and Chris Chiego, who taught me how to think victoriously.

To Drs. Leonard McCoy and Benjamin Franklin Pierce, who taught me how to think medically.

And finally, to Linna Duan, who taught me to stop thinking all the time.

## ABSTRACT

A marvel of evolution, the adaptive immune system has the capacity to respond to almost any foreign antigen in a highly specific manner. Antibodies, Y-shaped glycoproteins containing both diverse variable regions responsible for antigen binding and constant regions responsible for effector function, are a key part of this capacity. However, this vast diversity comes with several drawbacks, one of which is the fact that the immune system can deleteriously respond to self-antigens. The focus of this thesis is to characterize the role of class-switching (the changing of antibody constant regions) in the pathogenesis of autoimmune disease, and in particular to trace the lineage of antigen-specific autoreactive B cells by analyzing clonal relationships between antibodies of different constant regions. Analyzing such lineages has the potential to shed light on mechanisms of autoantibody-mediated disease pathogenesis, leading to better understanding of autoimmunity and better therapeutics.

The work presented in this thesis focuses on pemphigus vulgaris, or PV, a model antibody-mediated autoimmune disease characterized by a response to the cell adhesion protein desmoglein (Dsg) 3, which holds keratinocytes together in the epidermis. An enigmatic feature of this disease is the predominance of antibodies from the IgG4 subclass during active disease, which ordinarily appears to have few effector functions and may serve as a “brake” on the immune system in the setting of continuous stimulation by antigen. PV patients also display autoantibodies of the IgG1 subclass during disease and remission, but the relationship between IgG1 and IgG4 in the disease is unclear. Because the majority of cases of PV also harbor anti-Dsg

antibodies of the IgA1 and IgA2 subclasses, we sought to determine the relationships between autoantibodies belonging to each of these subclasses. First, we address whether the same anti-Dsg variable region, grafted onto either IgG1 or IgG4 constant regions, can show differing affinity or pathogenicity, in order to determine whether antibody subclass is directly modulating pathogenic effect (chapter 2). Finding that the subclass has very little effect on antibody affinity, pathogenicity, or epitope preference, we then sought to determine whether B cells expressing autoantibodies of different subclasses share lineages, indicating common pathways of development (chapter 3). Using a combination of antigen-specific antibody cloning through phage display, and next-generation sequencing of subclass specific repertoires in a panel of PV patients, we managed to trace 80 lineages of anti-Dsg B cells across all four subclasses tested. In particular, we found that anti-Dsg IgG4 B cells, which are believed to be central to disease pathogenesis, tended to not share lineages with other subclasses, and in general do not appear to share a precursor-product relationship with anti-Dsg IgG1 B cells. We have also found that anti-Dsg IgA1 and IgA2 were tightly related and often arose directly from IgG precursors. These findings are key to understanding the role of class-switching in the pathogenesis of PV, and may shed light on the class-switch mechanisms driving other autoimmune diseases and states of chronic antigen stimulation.

## PEER REVIEWED PUBLICATIONS

1. Ellebrecht CT‡, **Mukherjee EM‡**, Zheng Q, Choi EJ, Reddy SM, Mao X, Payne AS. Lineage analysis of immune repertoires in pemphigus vulgaris reveals distinct pathways for development of autoreactive IgG4 versus IgG1 and IgA B cells. Submitted to Cell Reports

This work is the basis of Chapter 3 of this thesis

2. Lo AS‡, Mao X‡, **Mukherjee EM‡**, Ellebrecht CT, Yu X, Posner MR, Payne AS, Cavacini LA. (2016) Pathogenicity and Epitope Characteristics Do Not Differ in IgG Subclass-Switched Anti-Desmoglein 3 IgG1 and IgG4 Autoantibodies in Pemphigus Vulgaris. PLoS ONE 11(6): e0156800. doi:10.1371/journal.pone.0156800

This work is the basis of Chapter 2 of this thesis

‡Indicates co-first authorship



## ATTRIBUTIONS

### INTRODUCTION:

-Figure 1-5 is adapted from:

Berkowska MA, Driessen GJA, Bikos V, Grosserichter-Wagener C, Stamatopoulos K, Cerutti A, et al. Human memory B cells originate from three distinct germinal center-dependent and -independent maturation pathways. *Blood*. 2011 Aug 25;118(8):2150–8.

### CHAPTER 2 (contains data from publication 2 on previous page):

- Agnes Lo produced F779 and F706 IgG1 and IgG4, and conducted ELISA on SCC cells
- Xuming Mao conducted internalization assays and CHO cell binding assay
- Eun Jung Choi conducted epitope mapping experiments

### CHAPTER 3 (contains data from publication 1 on previous page)

- Eun Jung Choi and Shantanu Reddy assisted with screening APD libraries and characterization of clones
- Christoph Ellebrecht produced the NGS libraries and the IgA phage display libraries
- Eun Jung Choi conducted epitope mapping experiments
- Qi Zheng conducted the sequencing pipeline and provided sage bioinformatics advice, until the fine lineage analysis using ChangeO
- Namita Gupta from the lab of Steven Kleinstein (Yale) offered much useful advice on the use of Change-O
- Ike Jose (NAQT) and Chris Chiego (Penn) advised extensively on R and Python coding

## TABLE OF CONTENTS

TITLE PAGE	i
ACKNOWLEDGEMENTS	ii
ABSTRACT	iii
PEER-REVIEWED PUBLICATIONS	v
ATTRIBUTIONS	vi
TABLE OF CONTENTS	vii
LIST OF TABLES	ix
LIST OF FIGURES	x
LIST OF ABBREVIATIONS	xi
<b>CHAPTER 1: INTRODUCTION</b>	<b>1</b>
1.1 Antibody Development	2
1.2 Pemphigus as a Model	18
1.3 Methods and Challenges of Antibody Repertoire Sequencing	26
1.4 Experimental Approach	40
1.5 Figures and Tables	46
Table 1-1. Summary of Pemphigus and Its Subtypes	46
Figure 1-1. VDJ Recombination at Heavy Chain Locus	47
Figure 1-2. VJ Recombination at Light Chain Locus	48
Figure 1-3. CDRs and FWRs are Interspersed in the Variable Region	49
Figure 1-4. Sample Clonality Plot	50
Figure 1-5. Human Heavy Chain Constant Region Locus	51
<b>CHAPTER 2: Swapping the Constant Region in Pathogenic IgG1 and IgG4 Autoantibodies in Pemphigus Vulgaris Does Not Alter Blister-Inducing Ability or Epitope Preference</b>	<b>52</b>
2.1 Abstract	52
2.2 Introduction	53
2.3 Materials and Methods	54
2.4 Results	60
2.5 Discussion	62
2.6 Figures and Tables	65
Figure 2-1. IgG1 and IgG4 PV mAbs Target Epitopes Within the Same Dsg3 Extracellular Domain	65
Figure 2-2. Immunoreactivity of anti-Dsg3 PV mAbs by ELISA and IIF	66
Figure 2-3. IgG1 and IgG4 Variants of anti-Dsg3 PV mAbs are Pathogenic and cause Suprabasal Blisters	67
Figure 2-4. Pathogenic F706 and F779 IgG1 and IgG4 Cause Loss of Cell Surface Dsg3	68

<b>CHAPTER 3: Lineage analysis of immune repertoires in pemphigus vulgaris reveals distinct pathways for development of autoreactive IgG4 versus IgG1 and IgA B cells</b>	69
3.1 Abstract	69
3.2 Introduction	70
3.3 Materials and Methods	72
3.4 Results	82
3.5 Discussion	92
3.6 Figures and Tables	100
Figure 3-1. Schematic of NGS Library Amplicon	100
Figure 3-2. Deriving CDR3 Bounds for Clonality	101
Figure 3-3. Autoreactive Lineages Bridge Subclasses	102
Figure 3-4. Connectivity of the Dsg-Reactive and Global Repertoire	103
Figure 3-5. Selected Trees from IgG4-Containing Multi-Subclass Lineages	104
Figure 3-6. Lineage Tree of PV16.G4.D333.6	105
Figure 3-7. Lineage Tree of PV8.G1.D1.P3.6	106
Figure 3-8. Lineage Tree of PV16.A2.D3.P3.10	107
Figure 3-9. Lineage Tree of PV16.G4.D3131.16	108
Figure 3-10. Lineage Tree of PV16.A1.D1.P2.7	109
Figure 3-11. Lineage Tree of PV16.G1.P2.14	110
Figure 3-12. Lineage Tree of PV16.G1.D313.10	111
Figure 3-13. Lineage Tree of PV8.A1.D3.3.2	112
Figure 3-14. Lineage Tree of PV8.A1.D1.P3.2	113
Figure 3-15. Lineage Tree of PV8.A1.D3.3.26	114
Figure 3-16. Lineage Tree of PV8.A1.D3.3.15	115
Figure 3-17. Lineage Tree of PV16.A1.D1.P2.8	116
Figure 3-18. Lineage Tree of PV17.A1.D1.P4.6	117
Figure 3-19. Lineage Tree of PV18.A1R.D3.P4.20	118
Figure 3-20. Lineage Tree of PV17.A1.D1.P3.10	119
Figure 3-21. ELISA Reactivity of Cloned Lineage Intermediates	120
Figure 3-22. Quantification of Sampling Coverage by Asymptotic Diversity Estimation	121
Figure 3-23. Error Correction for Lineage Tree Production	122
Table 3-1. Patient Characteristics	123
Table 3-2. Index of Lineage Trees	124
Table 3-3. Sequencing Metadata	125
Table 3-4. Anti-Dsg Clones Isolated by Antibody Phage Display	126
Table 3-5. Capture-Recapture Analysis Across Biological and Technical Replicates	127
Table 3-6. Summary of Epitope Mapping Experiments	128
<b>CHAPTER 4: DISCUSSION AND FUTURE DIRECTIONS</b>	129
4.1 Summary of Findings	129
4.2 Discussion and Future Directions	131
4.3 Conclusion	144
<b>CHAPTER 5: REFERENCES</b>	146

## LIST OF TABLES

Table 1-1. Summary of Pemphigus and Its Subtypes	43
Table 3-1. Patient Characteristics	118
Table 3-2. Index of Lineage Trees	119
Table 3-3. Sequencing Metadata	120
Table 3-4. Anti-Dsg Clones Isolated by Antibody Phage Display	121
Table 3-5. Capture-Recapture Analysis Across Biological and Technical Replicates	122
Table 3-6. Summary of Epitope Mapping Experiments	123

## LIST OF FIGURES

Figure 1-1. VDJ Recombination at Heavy Chain Locus	44
Figure 1-2. VJ Recombination at Light Chain Locus	45
Figure 1-3. CDRs and FWRs are Interspersed in the Variable Region	46
Figure 1-4. Sample Clonality Plot	47
Figure 1-5. Human Heavy Chain Constant Region Locus	48
Figure 2-1. IgG1 and IgG4 PV mAbs Target Epitopes Within the Same Dsg3 Extracellular Domain	61
Figure 2-2. Immunoreactivity of anti-Dsg3 PV mAbs by ELISA and IIF	62
Figure 2-3. IgG1 and IgG4 Variants of anti-Dsg3 PV mAbs are Pathogenic and cause Suprabasal Blisters	63
Figure 2-4. Pathogenic F706 and F779 IgG1 and IgG4 Cause Loss of Cell Surface Dsg3	64
Figure 3-1. Schematic of NGS Library Amplicon	95
Figure 3-2. Deriving CDR3 Bounds for Clonality	96
Figure 3-3. Autoreactive Lineages Bridge Subclasses	97
Figure 3-4. Connectivity of the Dsg-Reactive and Global Repertoire	98
Figure 3-5. Selected Trees from IgG4-Containing Multi-Subclass Lineages	99
Figure 3-6. Lineage Tree of PV16.G4.D333.6	100
Figure 3-7. Lineage Tree of PV8.G1.D1.P3.6	101
Figure 3-8. Lineage Tree of PV16.A2.D3.P3.10	102
Figure 3-9. Lineage Tree of PV16.G4.D3131.16	103
Figure 3-10. Lineage Tree of PV16.A1.D1.P2.7	104
Figure 3-11. Lineage Tree of PV16.G1.P2.14	105
Figure 3-12. Lineage Tree of PV16.G1.D313.10	106
Figure 3-13. Lineage Tree of PV8.A1.D3.3.2	107
Figure 3-14. Lineage Tree of PV8.A1.D1.P3.2	108
Figure 3-15. Lineage Tree of PV8.A1.D3.3.26	109
Figure 3-16. Lineage Tree of PV8.A1.D3.3.15	110
Figure 3-17. Lineage Tree of PV16.A1.D1.P2.8	111
Figure 3-18. Lineage Tree of PV17.A1.D1.P4.6	112
Figure 3-19. Lineage Tree of PV18.A1R.D3.P4.20	113
Figure 3-20. Lineage Tree of PV17.A1.D1.P3.10	114
Figure 3-21. ELISA Reactivity of Cloned Lineage Intermediates	115
Figure 3-22. Quantification of Sampling Coverage by Asymptotic Diversity Estimation	116
Figure 3-23. Error Correction for Lineage Tree Production	117

## LIST OF ABBREVIATIONS

PV – Pemphigus Vulgaris  
PF – Pemphigus Foliaceus  
MRCA – Most Recent Common Ancestor  
GL – Germline  
BCR – B-cell Receptor  
ELISA – Enzyme-Linked Immunosorbent Assay  
IIF – Indirect Immunofluorescence  
DIF – Direct Immunofluorescence  
mAb – Monoclonal antibody  
scFv – Single-chain variable fragment  
Dsg – Desmoglein

# **CHAPTER 1:**

## **Introduction**

Adaptive immunity in higher vertebrates, like humans, relies on generating highly diverse antibody molecules capable of responding to almost any target. While conferring protection against a variety of antigens, the adaptive immune system also has the power to respond to self-antigens, causing autoimmunity. As a geneticist, I view the adaptive immune system as a Darwinian mechanism, searching through a staggeringly diverse space of biological sequences. This can be analogized as a group of evolving and mutating B cells walking through an antibody-encoding Library of Babel\*, finding useful, protective antibody sequences amongst a vast background of irrelevant and potentially dangerous ones. The focus of my thesis work has been understanding the development, lineage, and sequence properties of autoreactive B cells in the model antibody-mediated autoimmune disease pemphigus vulgaris, to better understand the pathogenesis of autoimmunity. Separating the features of autoreactive cells from the larger repertoire will not only shed light on the pathogenesis of autoimmune disease, it may also help design better treatment strategies for the ever-increasing population of autoimmune disease sufferers worldwide (5).

---

\**The Library of Babel* is a short story by the Argentinian author Jorge Luis Borges(1). It considers a fictional universe containing all possible 410-page books housed in an “indefinite, perhaps infinite” number of linked hexagonal rooms. A similar analogy has been used to describe the total space of DNA or protein sequences by the philosopher Daniel Dennett (calling it the “Library of Mendel”)(2) and others (3)(4).

## **1.1 Antibody Development**

### **1.1.1 Introduction**

Antibodies, or immunoglobulins (Ig), are highly diverse glycoproteins generated as a key component of the adaptive immune system in higher vertebrates (6). The antibody macromolecule consists of four polypeptide chains, with two heavy and two light chains linked together by disulfide bonds forming a Y shape. The variable region of each monomer is contributed to by both the heavy and light chain, and serves to recognize and bind to antigen. The constant region, by contrast, determines the isotype and effector function of each antibody. The genesis, development and genetic basis of antibody constant and variable region diversity will be described in this section.

### **1.1.2 Antibody Genesis**

In humans, the peptides that comprise a mature antibody are encoded by separate Ig loci located on chromosome 14 (heavy chain, IgH), chromosome 2 (kappa light chain, IgK), and chromosome 22 (lambda light chain, IgL)(7)(8)(9). During development, each B-cell precursor rearranges these loci in a stepwise process called V(D)J recombination, so named because each Ig heavy chain locus contains several variable (V), diversity (D), and joining (J) segments that will be joined together to generate a contiguous immunoglobulin-coding DNA segment (10)(11). The light chain loci have a similar architecture and are rearranged in a similar fashion, though without



the D genes; therefore, the final heavy chain of an antibody is coded by three apposed gene segments, while the light chain is coded by two.

Hematopoietic stem cells in the bone marrow differentiate into multipotent progenitors, then common lymphoid progenitors (CLP), then pro-B cells, which are committed to the B cell lineage via the expression of the transcription factor Pax5 (12). At the pro-B cell stage, the Ig loci have not yet been made productive. At this point, transcription factors including BCL11A/B, FOG1, and FOG2 induce expression of the heterodimeric RAG1/2 recombinase, which initiates the process of VDJ recombination at the heavy chain locus (13)(14)(15)(16). First, a D gene and J gene are randomly selected, and RAG1/2 creates a double-strand break in the recombination signal sequences (RSS) adjacent to the selected D and J exons; the intervening segment is ligated to form a circular intermediate called the signal end complex, and the sequence adjacent to the D and J genes self-anneal to form hairpin structures (17)(18)(19). At this stage, the Artemis nuclease nicks and opens each hairpin, which generates a linear stretch of DNA containing palindromic (P) nucleotides (so-called since hairpins necessarily have palindromic sequences to self-anneal). This stretch of P nucleotides can subsequently be chewed back by variety of exonucleases to chew back the DNA, and/or extended by terminal deoxynucleotidyl transferase (TdT), which adds non-template (N) nucleotides adjacent to the original P nucleotides to generate additional diversity (20)(21)(22)(23)(24)(25). Subsequently, these two D and J genes are recombined together in an error-prone process of non-homologous end-joining (NHEJ)(26)(27). The

same process then occurs between a selected V gene and the DJ segment, creating one continuous VDJ region encoding the heavy chain variable region (Figure 1-1) (28).

At this stage, the mature heavy chain with an IgM constant region, paired with a surrogate light chain, is expressed on the surface of the developing B cell as part of the pre-B cell receptor (pre-BCR), along with invariant Ig-alpha and Ig-beta signaling molecules (29). Signaling through this receptor, which is only possible when the product of the heavy chain locus can successfully pair with the surrogate light chain, ends VDJ recombination at the IgH locus by down-regulating RAG1/2 (30)(31)(32). Furthermore, signaling at this step ensures that only one antibody heavy chain is expressed per B cell, in a process called allelic exclusion (33)(34)(35). This signaling allows the pro-B cell to transition to a pre-B cell, which subsequently divides repeatedly and rearranges its light chain locus in a similar fashion (Figure 1-2). In this stage, the pre-B cell does not yet express mature IgM on its surface, but does show cytoplasmic IgM heavy chains, which are rapidly degraded as they cannot yet pair with a mature light chain (36). After the light chain rearranges and pairs properly with the already-rearranged heavy chain, the pre-B cell becomes an immature B-cell and expresses a mature BCR complex on its surface (37). After crossing a few critical checkpoints detailed in section 1.1.5, immature B-cells move to the periphery, where they can divide and diversify to give rise to clonal lineages originating from a single VDJ recombination event. These later steps of antibody maturation and diversification in the periphery are discussed in later sections of this chapter.

Note that at the point of the B cell exiting the bone marrow, there are three key sources of sequence diversity in the variable heavy chain. First, the combinatorial diversity generated by selecting different V, D and J genes is substantial. According to the IMGT database, in the human IgH locus, there are 38-46\* functional V genes, 23 functional D genes, and 6 J genes, generating 4002-5244 different joins (39)(40)(41). Human light chains are encoded by 29-33 V and 4-5 J chains in the lambda locus or 31-35 V and 5 J chains at the kappa locus, for a total of 271-340 different joins (41)(42)(43)(44). Random pairing of heavy and light chains generates a final combinatorial diversity of between  $1.08 \times 10^6$  and  $1.78 \times 10^6$ . This is further increased by the imprecise nature of each V-D and D-J join in the heavy chain and V-J joins in the light chain, resulting in a final naïve B-cell repertoire diversity of approximately  $10^{11}$  in a single individual and theoretically greater than  $10^{14}$  across the entire space of variable region sequences (45)(46)(47).

However, this diversity comes at a significant cost. The NHEJ mechanism doesn't take reading frame into account, meaning that approximately two-thirds of recombination events will not produce functional protein (48)(49). When these nonproductive rearrangements fail to signal through the pre-BCR complex, they can be rescued by subsequent rearrangements on the same chromosome (selecting V genes upstream or J genes downstream of the unproductive VDJ locus) or, if those options are

---

\* IMGT reports a range, because there's variability from person to person in the number of functional V, D, and J genes in each locus. The loci also contain non-functional ORFs (which have an open reading frame but lack other features needed for function) and pseudogenes (which contain stop codons or frameshifts). For more information on the variability in the germline VDJ complement, please see (38)

exhausted, by subsequent rearrangement on the other chromosome. A similar process occurs in the light chain locus at the pre-B cell stage, resulting in a productive V-J rearrangement from either the lambda or kappa locus and isotypic exclusion of the other (35)(50). In sum, this results in approximately half of all pro-B cells going on to produce functional antibody (41).

The sequence structure of a mature antibody variable region gives some insight into its function (Figure 1-3). Each heavy and light chain can be divided into three loop-shaped hypervariable complementarity-determining regions (CDRs) interspersed with relatively rigid beta sheets called framework regions (FWRs) (51)(52). In general, CDRs are responsible for the antigen specificity of a given antibody, while the FWRs are responsible for stabilizing the antibody's three-dimensional structure, though this distinction isn't absolute (53)(54)(55)(56)(57)(58). CDR1 and CDR2 are entirely encoded by the V gene in both the heavy and light chain, and the CDR3 spans the V-J junction in the light chain and the V-D-J junction in the heavy chain. Because the CDR3 spans all segments and the junctional diversity therein, it is almost impossible for two independent pro-B cells to give rise to the same heavy chain CDR3 region after recombination, and therefore the CDR3 is commonly used as a marker of clonal relatedness between B cells (59).

As the B cell circulates and migrates into secondary lymphoid organs, it undergoes further diversification through two processes, somatic hypermutation and class switching. These processes, and their effects on the repertoire, will be described in the following sections.

### 1.1.3 Variable Region Development

Immature B cells further differentiate into transitional B cells as they exit the bone marrow (60)(61). At this stage, the transitional B cell migrates to follicles in secondary lymphoid organs (lymph node, spleen, Peyer's patches in the ileum, etc.). Either during or after its migration, the transitional B cell further differentiates into a mature naïve B cell, capable of responding to antigen and co-stimulatory signals (61)(62). At this stage, the B cell can enter a germinal center (GC) within a secondary lymphoid organ, which are sites of rapid B cell proliferation and differentiation (63). Before entering the GC, B cells can bind to and endocytose circulating antigen using its BCR, subsequently presenting linear peptides derived from said antigen on their surface using major histocompatibility complex (MHC) (64)(65). At this point, the antigen-experienced B cell can enter the GC, where it can form an immunological synapse with a helper T cell ( $T_H$  cell) expressing a cognate T-cell receptor (TCR) for the same antigen\*\* (67)(68). Through this MHC:TCR interaction, as well as co-stimulatory ligand-receptor interactions like CD40:CD40L, CD80:CD28, and CD86:CD28, the antigen-experienced B cell undergoes clonal expansion and further maturation, fine-tuning its affinity for its target antigen (69)(70)(71)(72).

The key molecule that catalyzes B cell affinity maturation is Activation-Induced Cytidine Deaminase (AID), a tightly-controlled enzyme which catalyzes C-to-U deamination in the IgH, IgL, and/or IgK variable regions in a transcription-dependent

---

\*\* TCRs undergo a similar VDJ recombination process as antibodies, though their function is to respond to linear peptides displayed on MHC molecules by antigen-presenting cells. The theoretical maximum of TCR diversity approaches  $10^{20}$ , though humans only have  $\sim 10^{12}$  total T cells. For more information, see (66)

manner (73)(74)(75)(76). The subsequent G:U mismatch can be dealt with after MSH2/MSH6 binding by the mismatch repair, or by uracil-DNA glycosylase creating an abasic site that is subsequently filled by base excision repair (77)(78)(79)(80). The former pathway can also induce mutations in nearby A/T residues on the same strand by recruiting the error-prone DNA polymerase eta (81)(82). This process, called somatic hypermutation (SHM), creates point mutations across the variable region, substantially increasing the total sequence space available to the immune system for antigen binding and antigen-driven selection. Mutations that increase the affinity of the BCR are subsequently selected and propagated by outcompeting lower-affinity sequences for limited T cell help, resulting in a clonally-related B cell pool with greatly increased affinity for the target antigen (83). These affinity-matured B cells can undergo further differentiation into plasma cells and memory cells (84). The former take up long-term residence in specified niches in the bone marrow, mucosa-associated lymphoid tissue (MALT), and spleen and secrete soluble antibody into the bloodstream and lymphatic system like so much biochemical artillery, while the latter express matured antibody in a transmembrane form, lying in wait to respond when re-challenged with the same antigen (85)(86)(87).

There also exists an alternative pathway for B cell maturation outside the GC, called the “extrafollicular pathway”, where B cells can mature without T cell help. This pathway is particularly important in first-line defenses requiring rapid production of antibody, and appears to function as part of the innate immune system. For example, in the marginal zone of the spleen, which samples antigens from circulation, simultaneous

engagement of toll-like receptors (TLRs) and BCR on the B cell surface can serve as a co-stimulatory signal, allowing B cell activation and differentiation into antibody-secreting plasma cells in the absence of T cells (88)(89). Antigen-presenting dendritic cells, which reside in areas in contact with the environment, can also directly induce activation and class switching by engaging BCR and secreting the B cell activating factors BLyS and APRIL (90). Finally, intestinal epithelial cells can sense bacterial antigens in the gut through TLRs, secreting APRIL and activating dendritic cells to induce T-independent class switch to IgA (91). In general, these extrafollicular responses create low-affinity antibodies with fewer somatic mutations than in B cells that pass through the GC, making these two repertoires, in theory, distinguishable (92,93).

Mutations induced during SHM are not perfectly uniform across the variable region, for a variety of reasons. AID preferentially mutates cytosines belonging to the “hotspot” motifs WRCY/RGYW and avoids cytosines in the SYC/GRS “coldspot” motif (94)(95)(96)(97). A secondary SHM hotspot, WA/IW, is due to DNA polymerase eta activity (98)(99)(100). High-throughput sequencing work has largely corroborated this hotspot-coldspot model, with the additional complication that hotspots can show a wide range of “hotness”, i.e. mutability (101). It has also been shown that through evolutionary pressure, hotspot motifs, and in particular hotspot motifs in which the mutable cytosine is in a position likely to result in amino acid replacement upon mutation, have been preferentially clustered into the CDR regions, while coldspot motifs segregate to FWR regions (102)(103)(104)(105). In theory, this ensures that SHM

preferentially mutates sections of the variable region responsible for antigen-binding, and avoids regions required for structural integrity.

In addition to point mutations, SHM can also add insertions and deletions (indels) to the variable region (106). Up to 6.5% of circulating B-cells encode indel-containing antibody variable regions; though rare, they add substantially to the overall diversity of the antibody repertoire (107). Indels tend to occur near AID hotspots, tend to localize to CDRs or CDR-adjacent residues, and rapidly decrease in frequency with length (108)(109). While the exact mechanism of indel creation isn't known, in vitro studies have shown that AID is necessary for this mechanism of diversification (107), and in vivo studies have shown that indels preferentially occur in repetitive or loop-forming tracts of DNA near AID hotspots, suggesting a model in which a repair polymerase "slips" while correcting an AID-induced mutation, either creating a deletion or duplicating adjacent nucleotides to create an insertion (110).

SHM-induced indels allow the adaptive immune system to sample relatively inaccessible regions of sequence space; those B cells fortunate enough to reach these obscure rooms in the library can produce antibodies with particularly interesting reactivity. For example, an antibody capable of neutralizing both the 1918 and 2009 H1N1 influenza virus shows a 3-amino acid insertion in the heavy chain FR3 region; crystallization showed that this insertion pushed the CDR2 loop into a useful conformation for binding to hemagglutinin and was necessary for the antibody's ability to neutralize both strains (111–113). In another study, around 40% of surveyed broadly-neutralizing HIV antibodies (bnAbs) were found to have variable region indels, and in



several cases the indels are either required for high-affinity binding or required for subsequent affinity maturation of the bnAb lineage (114). In interesting recent work, two patients from a malaria-endemic region in Kenya were found to express unusual broadly-reactive antibodies against *Plasmodium falciparum*. Shockingly, these antibodies contained not mere SHM-derived indels, but a 100-amino-acid insertion between the V and D genes consisting of the entire collagen-binding domain of the collagen receptor LAIR1. The mechanism for this last example is remains unknown - authors found that the LAIR1 insert contained cryptic RSS site on either side, but both copies of LAIR1 were still extant in the genomic DNA, which is inconsistent with RAG1's copy-paste mechanism (115). These results all suggest that indels are a particularly powerful way of diversifying the antibody repertoire, with the potential of creating reactivity to traditionally "difficult" antigens.

The combination of a "seed" VDJ recombination in the bone marrow and the Darwinian pull of SHM in the periphery allows a single unique variable region in a pro-B cell to give rise to an entire lineage of B cells after affinity maturation. Conceptually, a particular B cell lineage can be discerned by its usage of shared V, D and J genes and CDR3 sequence (with some mutation within the CDR3), with the highly diverse CDR3 sequence serving as a natural barcode propagated within said lineage. Therefore, the antibody repertoire can be viewed as a series of disjoint lineages, each originating from a single VDJ recombination event. This idea is reflected by high-throughput sequencing experiments. In these data sets, binning all sequences by their assigned V and J gene (which are presumed shared by any given lineage), then calculating the pairwise

nearest-neighbor distance between CDR3 sequences within a given bin, gives a bimodal distribution with a “near” peak corresponding to pairs of clonally-related sequences and a more pronounced “far” peak corresponding to unrelated CDR3 (Figure 1-4) (116). The task of sorting the repertoire into lineages, and the general difficulties of antibody repertoire sequencing, will be discussed in the final section of this introduction.

#### **1.1.4 Isotypes and Class-Switching**

The constant region of an antibody is encoded by the C-terminal portions of the heavy and light chain. The heavy chain constant region, in particular, is the sole determinant of the antibody isotype, which determines the antibody’s effector function (117). In humans, there are 5 antibody isotypes (IgM, IgD, IgG, IgA, and IgE), some of which can be further divided into subclasses (IgG1 through IgG4, and IgA1 through IgA2)(118). The 3’ end of the heavy chain locus contains exons corresponding to each of these isotypes and subclasses, with a small repetitive switch region (S-region) and intronic promoter (I-region) immediately 5’ of each exon (Figure 1-5) (119)(120).

Late-stage transitional B cells display both membrane-bound IgD and membrane-bound IgM through alternative splicing (121)(122)(123)(124)(125). Later, during affinity maturation in the GC, AID (in addition to catalyzing SHM) can also catalyze class-switch recombination (CSR) (126)(127)(128). In this process, the cytokine milieu generated by T<sub>H</sub> cells and other cells within the GC epigenetically triggers transcription from a donor and acceptor I-promoter in the IgH locus. The single-stranded DNA at the transcription bubbles serves as a substrate for AID, inducing double-strand breaks due to a very high

density of overlapping AID hotspot motifs. These lesions are subsequently repaired by NHEJ between corresponding donor and accept S-regions, “looping out” the DNA between them and forming a “switch circle” byproduct (129).

CSR is therefore a one-directional process – a B cell can switch from a 5' exon (like IgM) to a more 3' one (like IgE), but the reverse is impossible. In particular, it is worth noting that class-switching is often sequential, rather than direct from IgM. For example, by cloning switch circles from ex vivo allergen-challenged human B cells, one group showed that B cells can switch directly from IgM to IgE, or switch sequentially from IgM to IgG4 to IgE (130). Similar results have been found in allergen-challenged mice, with one study noting that a sequential switch was necessary to develop high-affinity IgE (131)(132). A high throughput sequencing study of IgE lineages from the peripheral blood of 33 human subjects showed, based on shared somatic mutations between sequences of different subclasses, that IgE-expressing B cells predominantly arose through sequential switch, primarily from somatically-mutated clones in IgG1 (133). Another group proposed a model of sequential class-switching between IgG subclasses (IgG3 > IgG1 > IgG2 > IgG4) over the course of an immune response, which is supported by the presence of the requisite switch remnants in IgG1 and IgG2 B cells (134)(135)(136). In a more expansive study, Stephen Quake's group has recently published a comprehensive landscape of human CSR by sequencing peripheral blood cells from 11 pairs of healthy human twins and lineage tracing every variable region sequence found. This work demonstrated that direct switching from IgM to downstream

subclasses like IgA2, IgG4, and IgE is rare, and the majority of switches to these subclasses pass through more upstream subclasses like IgG1 or IgA1 first (137).

Much as SHM diversifies the antigen-binding function of a given antibody, CSR diversifies the effector function of each antibody (though with ten orders of magnitude less diversity). A complete survey of each subclass and the signaling pathways leading to their synthesis is the subject of ongoing research and is beyond the scope of this thesis; however, those subclasses and cytokines most important for pemphigus pathogenesis will be reviewed in section 1.2.

### **1.1.5 Autoreactivity and Tolerance**

The vast capacity of the antibody repertoire to respond to nearly any antigen over the course of an organism's lifetime has a serious drawback – it has the capacity to produce antibodies against self-antigens (138)(139). A seminal study showed that a staggering 50 to 75% of immature B cells in the bone marrow react to cytoplasmic and/or nuclear antigens in HEp-2 cell-line extracts (a standard assay for autoreactivity) (140). There are several checkpoints in place to prevent these autoreactive clones from propagating, both within the bone marrow (central tolerance) and in the periphery (peripheral tolerance). These mechanisms are demonstrably impaired in several autoimmune diseases, including multiple sclerosis (141), rheumatoid arthritis (142), myasthenia gravis (143), Sjogren's syndrome (144), and systemic lupus erythematosus (SLE) (145), all of which demonstrate increased numbers of autoreactive B cell clones in

the peripheral blood. The mechanisms of these checkpoints will be elucidated in this section.

Upon expression of a pre-BCR at the pre-B cell stage, the nascent B cell is challenged by a panoply of self-antigens within the bone marrow, beginning the process of central tolerance (146). Subsequently, a high-affinity interaction between the pre-BCR and a self-antigen at this stage can trigger a number of different central tolerance mechanisms, including receptor editing, anergy, and clonal deletion (147)(148)(149). In receptor editing, a self-reactive pre-B cell arrests development, reactivates RAG1 and RAG2, and attempts to abolish self-reactivity by re-arranging the light chain locus (150). Typically, the kappa locus is rearranged first, then the lambda (151)(152)(153)(154). Subsequently, if the clone remains autoreactive, an analogous process can occur at the heavy chain variable region, replacing the current VH gene with a more upstream one (155)(156). This process can even occur serially, with multiple upstream VH genes tested for autoreactivity (157). However, the details remain controversial. Another study used a knock-in mouse model to show that that VH replacement occurred before light-chain rearrangement, suggesting that VH replacement isn't involved in editing autoreactive antibodies (158). Furthermore, while VH replacement typically results in lengthening of the CDR3 and leaves a "footprint" of the replaced VH gene that can theoretically be detected algorithmically in sequencing data (159)(160), other studies suggest many of these footprints may arise through other mechanisms, including sequencing error and chance (161). Therefore, while VH replacement is known to occur, both its overall frequency and its relevance to tolerance

remains controversial. Chapter 3 of this thesis includes analysis of cloned antibodies for autoreactivity, but for the above reasons it cannot reliably be determined which receptor editing processes these clones have undergone within the bone marrow before migration into the periphery.

If the B cell remains autoreactive after receptor editing, its differentiation is arrested and it undergoes apoptosis after 2-3 days in a process called clonal deletion (162)(163)(164). Deletion is the last central tolerance mechanism for eliminating a high-affinity autoreactive B cell before it enters the periphery; however, clones expressing BCRs with low affinity can still escape. At this stage, those low-affinity clones that react to abundant soluble antigen in the periphery can undergo tonic BCR stimulation in the absence of co-stimulatory signals; this causes them to down-regulate surface IgM, ignore mitogenic signals, and decreases their half-life, a state called anergy (165)(166)(167)(168)(169). Anergic mature naïve B cells are excluded from the GC and are unable to solicit T cell help; this prevents them from receiving critical survival signals from the follicular niche and results in their death (170)(171)(172). The importance of anergy in preventing autoimmunity can be seen in transgenic mice deficient for inhibitory components of the BCR, like SHP-1 and SHIP-1; both of these are required for anergy, and their knockout causes severe autoimmunity (173)(174)(175).

In healthy individuals, these central tolerance mechanisms manage to eliminate most autoreactive B cells, but 5-20% of mature naïve B cells remain autoreactive (140). At the level of GC and post-GC B cell development, there are a few more tolerance mechanisms to consider. As previously stated, T cell help is required for the production

of high-affinity antibodies (176), and more importantly a lack of T cell help in the GC leads to B cell apoptosis (177)(178). T cells must cross their own tolerance checkpoints, both within the thymus and the periphery; this decreases the chances that an escaping autoreactive B cell and its cognate T cell both survive and meet in the GC (179)(180).

Affinity maturation in the GC can create autoreactive clones, some even managing to differentiate into long-lived memory B cells (181)(182)(183)(184); however apoptosis usually prevents newly-autoreactive B cells from escaping the GC (185). These autoreactive B cells can undergo outright deletion in the periphery, as demonstrated in several double-transgenic mouse experiments in which an antigen and B cell expressing a cognate antigen-specific BCR are both expressed (186)(187). Subsequent experiments showed that in these systems, deletion occurs due to the absence of T cell help (188). Deletion isn't the only fate for a new autoreactive clone, however. There's some evidence that receptor editing acts on autoreactive clones in the periphery (189), though some mouse models suggest that heavy chain rearrangements in the periphery actually improve autoantigen binding, rather than abolishing it (190)(191). In other cases, further mutation and/or glycosylation of the variable region and further mutation "redeemed" autoreactive clones (192)(193).

A simple analogy can be drawn between autoimmunity and cancer. Just as cancer represents a failure of multiple checkpoints to prevent the unchecked cell division of initially rare aberrant cells, autoimmunity is the result of a (sometimes quite polyclonal) sub-population of autoreactive lymphocytes escaping multiple tolerance mechanisms. Just as understanding the mechanisms of tumorigenesis has ushered in

powerful targeted therapy for a variety of cancers, understanding the genesis of autoreactive B cells has the potential to herald a new era in targeted therapy for autoimmune disease. In chapter 3, we have combined antigen-specific screening and high-throughput sequencing to trace the lineage of autoreactive B cells across several subclasses, to better understand the origin and development of autoimmune disease. In particular, we're interested in tracing the B cell lineages responsible for the organ-specific autoimmune disease pemphigus vulgaris (PV), which we describe in the next section.

## **1.2 Pemphigus as a Model**

### **1.2.1 Introduction**

Pemphigus is a group of rare autoimmune diseases of the skin and mucous membranes characterized by serum antibodies to cell adhesion proteins, causing intraepidermal blistering (194)(195). The various subtypes of pemphigus are categorized based on the adhesion proteins targeted by autoantibodies, and show different clinical and histological presentations (Table 1-1). PV is the most common disease in this group, occurring in 0.1-0.5 of every 100,000 people (196). In particular, the incidence of PV seems to increase in populations with Middle Eastern, Jewish, or Mediterranean ancestry. For example, in Hartford County, Connecticut, the incidence of PV in the Jewish population was 3.2 per 100,000 people, while the overall adult population had an incidence of 0.42 per 100,000. In Isfahan, central Iran, the incidence is even higher, at 5



per 100,000 (197), though in the more southwestern city of Shiraz the incidence dropped to 0.59 per 100,000 in another, suggesting some degree of geographic variation (198). Among European populations, the UK seems to have a slight increase over the continent, with a PV incidence of 0.7 per 100,000 (199), compared to 0.5-1.5 per million in Switzerland (200), France (201), Finland (202), and Germany (203).

Clinically, PV is characterized by suprabasal blistering of the oropharynx and a positive ELISA against the cell adhesion protein desmoglein (Dsg) 3. Suprabasal blistering of the skin can occur with the subsequent development of antibodies against a homologous cell adhesion protein, Dsg1 (204)(205). This is reminiscent of another disease in the group, pemphigus foliaceus, which is characterized by antibodies to Dsg1 and superficial skin blistering. The serum autoantibody profile of PV is used to categorize it into its mucosal and mucocutaneous subtypes, the former expressing anti-Dsg3 antibodies, the latter expressing both anti-Dsg3 and anti-Dsg1. The reasons for this tissue distribution will be made clear later in this section.

Pemphigus is a model autoantibody-mediated disease for a few reasons. First, the autoantigens are well-characterized, with highly sensitive and specific ELISA assays available for both Dsg3 and Dsg1. Second, antibodies have been shown to be both necessary and sufficient for disease. For example, PV serum IgG causes blisters by passive transfer into neonatal mice, while IgG from the serum of healthy subjects and from PV serum with anti-Dsg3 antibodies removed by adsorption does not (206)(207). Furthermore, monoclonal anti-Dsg3 antibodies cloned from PV patients can cause blisters in neonatal mice and human skin (208)(209). These features make PV ideal for

studying an antigen-specific autoantibody response, and for understanding the evolution of an autoimmune repertoire.

### **1.2.2 Pathogenesis**

The association between pemphigus vulgaris and autoantibodies was first demonstrated in the 1960s, by immunofluorescence showing that patient sera binds in a cell-surface pattern to human skin (210)(211). Almost three decades later, the autoantigen was described by John Stanley's group as a novel epithelial cadherin, later cloned as Dsg3 (212)(213). Dsg3 mediates cell-cell adhesion as part of the desmosome, which connects keratinocytes in stratified squamous epithelia. Within the desmosome, desmogleins can undergo both homophilic interactions (214) and heterophilic interactions (215)(216). Intracellularly, the C-terminal domain of desmogleins bind to a complex of proteins including plakoglobin, plakophilin, and desmoplakin, which mediate interactions with the keratin cytoskeleton (217). This allows the desmosome to resist mechanical stress, such as rubbing of the skin; it is thus apropos that one of the clinical signs of PV is "Nikolsky's sign", in which rubbing of normal skin forms a blister (218).

The extracellular domain of Dsg3 consists four extracellular cadherin-like domains (EC1-EC4), which lie distal to a fifth extracellular domain dubbed EC5. The N-terminal EC1 and EC2 domains, which are the most and second-most distal to the cell surface, respectively, have been shown to mediate cis (e.g. interactions between proteins on the same cell) and trans interactions (interactions between proteins on different cells) (219)(220). Because of this, binding of autoantibodies to these N-

terminal domains can directly disrupt trans interactions by disrupting desmosomal assembly(221)(222) and/or causing internalization of Dsg3 into endosomes(223), leading to blistering.

The N-terminal preference of PV autoantibodies is well-established, though reactivity is not exclusive to that domain (224)(225). A longitudinal study of pemphigus sera from 53 patients showed predominant reactivity with the EC1 domain at all stages of disease, with decreasing percentages of sera showing reactivity as one moves to EC2-5 (226). A previous repertoire cloning experiment using phage display showed that antibodies capable of causing a blister in neonatal mice by injection (“pathogenic” antibodies) bound to the N-terminal 161 amino acids of Dsg3 (which covers EC1 and part of EC2), while non-pathogenic antibodies preferentially bound to more C-terminal domains (208). These studies all suggest that the autoantibodies most critical for PV pathogenesis bind to the N-terminal EC domains of Dsg3. However, the role of antibodies binding to C-terminal EC domains should not be overlooked. Injecting multiple non-pathogenic antibodies together can cause a blister, suggesting that so-called non-pathogenic antibodies can work synergistically to have pathogenic activity (227).

Dsg1 acts as a secondary target in PV, specifically appearing in its mucocutaneous variant (204,205). The anti-Dsg1 and anti-Dsg3 ELISA have been correlated to disease activity in skin and oral mucosa respectively (228), and anti-Dsg1 ELISA titer can predict skin relapse (229). These studies suggest that anti-Dsg1 antibodies are required for skin involvement in PV, and that anti-Dsg3 antibodies are

insufficient to cause cutaneous blistering. This feature has been explained by the fact that both Dsg3 and Dsg1 are expressed in the epidermis in an inverse fashion (Dsg3 being basal and Dsg1 superficial) (230)(231)(232), while Dsg3 is expressed much more strongly than Dsg1 across all layers of the oral mucosa (233). These observations led to the development of the desmoglein compensation hypothesis to explain pathogenesis. The hypothesis states that in the skin, Dsg1 is capable of mediating adhesion when Dsg3 function is abrogated because it is expressed throughout the epidermis (including the basal layer), while the expression of Dsg1 is both completely lacking in the basal layer and significantly less than that of Dsg3 throughout the mucosa, making it unable to maintain adhesion in the basal layer in the presence of anti-Dsg3 antibodies (234). Evidence for this theory mainly comes from experiments in transgenic mice. One experiment showed that antibodies to both Dsg1 and Dsg3 are required for blister formation in neonatal wild-type mice, while Dsg3-null neonatal mice only required antibodies to Dsg1 for blister formation (235). A transgenic mouse expressing Dsg3 ectopically in superficial layers of the skin is resistant to blistering when injected with anti-Dsg1 serum (236). And finally, Dsg3(-/-) mice rescued with a Dsg1 transgene are resistant to blistering upon injection with the anti-Dsg3 antibody AK23 (237).

From an antibody repertoire perspective, the fact that two homologous proteins are targeted by the same disease, in a successive fashion, is very interesting. It may be the case, for example that the anti-Dsg1 repertoire is entirely descended from anti-Dsg3 clones over the course of the autoimmune response – this would be a parsimonious explanation to explain the transition from mucosal to mucocutaneous PV. Supporting

this idea is the fact that antibodies cross-reactive to Dsg3 and Dsg1 have been isolated from PV patients (208). Alternately, the Dsg1 repertoire may arise completely independently from the Dsg3 repertoire, despite the homology between the two. This is supported by a study which showed that affinity purified anti-Dsg3 serum from PV patients lacks reactivity to Dsg1 and vice versa (238), though it leaves open the possibility that anti-Dsg1 and anti-Dsg3 antibodies share a lineage and somatic mutations have modulated their reactivity. In chapter 3 of this thesis, we analyze both anti-Dsg1 and anti-Dsg3 antibodies from a panel of patients through phage display as part of a larger effort to define lineage relationships among autoreactive B cell clones in PV.

### **1.2.3 The Features of Autoantibodies in Pemphigus Vulgaris**

Like a handful of other autoimmune diseases, like anti-MuSK Myasthenia Gravis (239) (240), thrombotic thrombocytopenia purpura (241), and others (242), PV shows an IgG4 predominance during active disease. In particular, PV patients in active stages of the disease predominantly show anti-Dsg3 IgG4 and less often anti-Dsg3 IgG1 (243). Furthermore, autoantibodies show a 4-8 fold enrichment in IgG4 vs IgG1 sera by ELISA (244). Other studies have demonstrated that healthy relatives of PV patients (rarely), HLA-matched controls (rarely), and patients in remission can show anti-Dsg3 antibodies of the IgG1 subclass, while those with active disease show both IgG1 and IgG4 anti-Dsg3 antibodies (245)(246). Another study suggested that during IVIG treatment, the levels of

anti-Dsg IgG4 correlated better with clinical disease activity than anti-Dsg IgG1, suggesting the IgG4 antibodies have a more critical role in mediating disease (247).

Structurally and functionally, the IgG4 subclass is unusual for several reasons. IgG4 appears to arise in situations of chronic antigen stimulation, including in beekeepers (248), chronic helminth infection (249)(250), and patients undergoing allergen desensitization therapy (251). Because the constant region of IgG4 cannot fix complement, only weakly binds to Fc receptors, and cannot form immune complexes, IgG4 antibodies are believed to interfere with the binding of other antibody subclasses to prevent overreaction of the immune system to a particular antigen (252)(253)(254)(255). Also intriguing is the reciprocal relationship between IgG1 and IgG4 over the course of an immune response. In beekeepers, repeated stimulation by bee venom shifts the immune response to bee antigens from an IgG1 to an IgG4 predominant state (256). In a form of pemphigus foliaceus endemic to Brazil (fogo selvagem), patients living in the endemic area in the preclinical stages have anti-Dsg1 IgG1, while patients with active disease show an IgG4 predominance (257). Taken collectively, these observations suggest that anti-Dsg autoantibodies undergo a class-switch from IgG1 to IgG4 during active disease, potentially acquiring mutations that cause epitope spreading or increase their affinity.

In addition to IgG1 and IgG4, the majority of PV patients have anti-Dsg IgA as well (243)(258)(259). Note that there exists a rare clinical entity called IgA pemphigus, which is characterized by a neutrophilic infiltrate and reactivity to desmocollins, but this is considered a separate disease (see Table 1-1) and is therefore not relevant to our

results. IgA is the subclass primary responsible for maintaining immunity on mucosal surfaces as a first line of defense and to maintain homeostasis with normal flora (260)(261). The mucosal immune system accomplishes this by using secretory IgA to continuously sample antigens from the surface of the mucosa and modifying its repertoire in response (262). There exist two IgA subclasses, IgA1 and IgA2; the former is more prominent in the serum, while the latter is more common in the GI and GU tracts (263)(264). Furthermore, because 25% of IgA secreting plasmablasts in the gut are polyreactive, the mucosal immune repertoire may be a harbor of autoreactivity (265). That fact, taken alongside the observation that PV often starts in the oral mucosa, led us to hypothesize that IgA (specifically, IgA1) may contain clones cross-reactive to Dsg3 that eventually switch to IgG4 and cause disease.

The coexistence of four different antibody subclasses, which normally have radically different functions in healthy subjects, all targeting 1-2 two autoantigens (Dsg1, Dsg3, or both) makes PV an ideal model disease for studying the role of class switching and antibody subclass in autoimmunity. An analysis of the lineage relationships between anti-Dsg IgG1, IgG4, IgA1, and IgA2 can be found in chapter 3 of this thesis. We now review a method of high-throughput B cell repertoire analysis using next-generation sequencing, which was used in chapter 3 to shed light on these questions.

## 1.3 Methods and Challenges of Antibody Repertoire Sequencing

### 1.3.1 Overview

The advent of high-throughput DNA and RNA sequencing has drastically changed the way that genomes and transcriptomes are studied. By isolating all of the DNA or RNA from a tissue sample (or even a single cell), scientists can gain unprecedented insights into the genome or transcriptome at a much finer resolution than allowed by previous microarray-based or Sanger-sequencing based efforts. Further improvements to the technology, like ultra-deep sequencing and sequencing of targeted loci (266), has allowed for detecting rare resistant viral sub-populations in HIV patients (267), rare subclonal populations in cancer (268)(269), and previously unknown splice variants of well-characterized genes (270).

High-throughput sequencing technology has also revolutionized the study of the B cell repertoire (46). For example, experiments sequencing bulk DNA or RNA from peripheral B cells have characterized the immune repertoire in healthy subjects (271)(272)(273), assayed changes in repertoire during aging (274), found convergent sequences across patients in infection (275), identified and lineage-traced broadly-neutralizing antibodies to HIV (276)(277), and studied the vaccination response (278)(279)(280)(281)(282). Repertoire sequencing of autoimmune are not as common yet, but experiments have shown exchange of pathogenic B cells between the periphery and cerebrospinal fluid in multiple sclerosis (283)(284)(285), analyzed clonal expansions over time in Sjogren's syndrome (286), and demonstrated polyclonal nature of lupus



flares (287). These studies have shed light on the fundamental mechanisms of autoimmunity and have the potential to guide the future development of therapy.

The problems of sequencing the antibody repertoire are many-fold, and push the limits of high-throughput sequencing technology in many ways. The most trivial of these is the overall length of the VDJ locus; at ~450 bp, reliable sequencing of the entire variable region reaches the upper limit of current sequencing read lengths, and including an identifiable stub of the constant region (20-40 bp) and accessory barcode sequences pushes this limit even further. Full length VDJ sequencing with acceptable read numbers has been recently made possible in the last few years by Illumina's MiSeq 2x300 bp paired-end platform, which outputs up to 25 million reads (288). Previously, experiments have either sacrificed coverage using primers within the FWR2 (271) or FWR3 (289) region, or sacrificed depth by using 454 Life Science's GS FLX Titanium platform, which touts read lengths up to 600-1000 bp but only obtains  $10^6$  reads per run (290)(291).

Other issues in repertoire sequencing are largely due to the fact that repertoire sizes push the  $10^{11}$  range. Each and every one of those 100 billion sequences may be important, in some broad sense, to any given antigen response. This has a few implications. First, detection and reliable annotation of rare sequences is vital, much like in ultra-deep sequencing; this is compounded by the fact that the B cell repertoire is spread amongst the entire blood, lymphatic system, and potentially other tissues, AND across several different B cell subtypes with differing functions. Furthermore, given the clonal evolution of the repertoire by SHM, the ability to distinguish closely-related

sequences from one another is critical to understanding the shape of a given antibody's lineage; this is especially difficult because amplification and sequencing error can confound such efforts. Next, unlike the small handful of SNPs or mutations found in a typical human or viral genome (which allows relatively painless annotation of sequencing reads against a reference), the human VDJ locus contains a large number of closely-related V, D and J genes, with many allelic variants, which are then further mutated by SHM (sometimes on the order of 40 or greater mutations from germline (292)), making VDJ assignment much more difficult. This is particularly difficult for the D gene, which lies in the middle of the hypervariable CDR3 region. Next, assigning given antibody sequences to clonal lineages is confounded by the above problem of VDJ assignment, by sequencing and amplification error, and by the simple fact that sequencing only captures the current VDJ sequence of a given B cell, and will not be able to interrogate its entire life story. Finally, the issue of assigning functional characteristics to a given antibody sequence is currently impossible computationally, requiring additional nontrivial experiments to be carried out alongside repertoire sequencing. Each of these challenges will be addressed in the subsequent sections.

### **1.3.2 Sampling and Reproducibility**

Humans have around 5 liters of blood, and a typical peripheral blood draw will capture only 0.1-1% of this. In addition, only 2% of the  $10^{11}$  B cells in the average human are circulating at any given time, with the rest sequestered in bone marrow, spleen, lymph nodes, mucosal tissue, etc (46). Both of these facts make sampling the overall

repertoire challenging. Early experiments performed on *Danio rerio* managed to evade both of these constraints by simply snap-freezing, lysing, and extracting the RNA from entire fish, making all  $3 \times 10^5$  B cells accessible to sequencing (293)(294). Repertoire sequencing experiments on mice also manage to avoid these constraints, because in addition to having a thousand-fold smaller B cell repertoire than humans, the lymphoid tissue of mice is more easily obtained by surgery (295)(296).

Because of these technical, practical, and ethical constraints, repertoire analysis in humans will necessarily be limited by sample. These limits can be assessed using biological and technical replicates for estimating the overall depth of sampling (47,296). Data from biological replicates (e.g. multiple blood or tissue samples from the same subject) quantifies the amount of biological undersampling, while technical replicates (e.g. multiple PCR or sequencing experiments on the same biological sample) quantifies the sufficiency of sequencing depth. This kind of replicate data is particularly useful for capture-recapture analysis, in which the degree of overlap (either in terms of lineages or individual sequences) between two samples is used to estimate the amount of the total population captured; this is similar to the mark-and-recapture analysis used to estimate populations in ecology (291)(293)(297). A popular unbiased statistic used for this purpose is Chapman's modification of the Lincoln-Petersen estimator, which uses the overlap between a pair of samples from a population to calculate the total population (297)(298).

A key assumption of capture-recapture methods is that each individual member of the population is equally likely to be "caught" by sampling (299); this assumption is

usually untenable, since lineage sizes and sequence copy numbers within the repertoire are not uniform, at either the RNA or DNA level. In other words, a B cell clone with multiple copies is more likely to be “trap happy”, e.g. more likely to be sampled multiple times, artificially inflating the overlap between samples. Methods to decrease the effects of trap heterogeneity are the subject of ongoing research, though none have yet been applied to antibody repertoire sequencing (300)(301).

Another ecology-inspired method, less subject to heterogeneity bias, is abundance-based rarefaction analysis, in which subsampling the reads from a given sample is used to create an accumulation curve (296)(302) . Some measure of diversity (e.g. number of unique sequences or lineages) in each subsample is plotted against the number of reads in the subsample, with the hope that the subsequent curve shows saturation as the number of subsampled reads approaches the actual number of reads. As the number of reads increases, the curves tend to quickly rise and plateau, as the common species (equivalent to large lineages) are quickly added to the curve, leaving only rare species (smaller lineages) to be found.

Recently, this analysis has extended this kind of analysis to include extrapolation and nonparametric asymptotic estimation of diversity, allowing the construction of smooth rarefaction curves using both interpolation and extrapolation via the Chao method (303)(304). The nonparametric nature of this analysis means its results are not reliant on the distribution of species in the population, unlike the capture-recapture estimation outlined above. The mathematical details are outside the scope of this thesis, however a thought experiment is illustrative. Imagine a sample containing one

thousand different lineages, each of them with only one read (singletons). Compare that with a sample containing 25 different lineages, 20 of which have 100 or more reads and 5 of which have a single read. It is easy to see why the second sample is closer to saturation than the first. If our sample predominantly consists of singletons, it suggests that collecting a larger sample will easily find new species. By contrast, if singletons are only a small proportion of the sample, further sampling will probably not find many new species. The Chao method takes the proportion of singletons, two-read species, etc up to 10-read species into account, using this to infer the number of unfound species for extrapolation.

A drawback of the Chao method is that it is exquisitely sensitive to the presence of singleton lineages (i.e. lineages with only one sequencing read assigned to them); this is a particular concern because sequencing error often inflates the number of low-frequency members in a sample (305)(306). Some non-statistical methods of removing erroneous sequences are discussed in the next section; however, for the purposes of diversity estimation, we can also statistically deflate the number of low-frequency counts according to some expected distribution. Sampling analysis in chapter 3 employs a method that uses Good-Turing frequency estimation (307)(308) to deflate the singleton count by generating an “expected” singleton count from the number of doubletons, tripletons, and quadrupletons (lineages containing 2, 3, and 4 reads, respectively), thus giving a more accurate estimation of diversity (309). However, it is worth noting that sequencing error inflates *all* low-frequency counts, including doubletons, tripletons, and quadrupletons, suggesting that even with statistical

correction we may be inflating our lineage counts slightly. Future developments in the field, including better statistical methods and technical improvements in sequencing, will lead to more accurate estimation of diversity and coverage in high-throughput sequencing datasets.

The effects of tissue choice on sampling are less well-known, as studies of B cell repertoires from non-blood tissues in humans has just begun. One group has conducted a repertoire sequencing study using RNA samples from peripheral blood, bone marrow, small intestine, lung, stomach, lymph node, tonsil, spleen, and thymus, discovering that the mucosal repertoire was quite different than that of the peripheral blood and lymphoid tissues in terms of VDJ use, CDR3 length, and somatic mutation frequency (310). Another group has also conducted B cell repertoire sequencing from both blood and cerebrospinal fluid in patients with multiple sclerosis, showing extensive exchange of B cells between the two compartments (311).

Because previous repertoire cloning studies in PV have used peripheral blood, managing to find several anti-Dsg clones (208), and because of the difficulty of sampling other compartments, the studies in chapter 3 are restricted to blood, and therefore may miss relevant B cell diversification events in the mucosa, spleen, or other compartments. It is entirely feasible, for example, that skin-draining lymph nodes in pemphigus contain distinct repertoire of anti-Dsg antibodies. Further studies comparing different tissue-specific repertoires in PV may shed light on their role in disease states.

### 1.3.3 Error Correction

In any high-throughput sequencing experiment, there are two main sources of mutational error – from amplification, and from sequencing. The Illumina MiSeq platform used in this thesis is known to be prone to substitution error, especially in GC-rich regions, on the order of 1-2 bases per 1000 nucleotides (312)(313). While this error rate may be sufficient for simple RNA-seq expression studies, it is unfortunately on the order of the mutation rate for SHM (which is approximately 1 in 1000 bp per cell division) (102). This is compounded by the fact that the mutation rate for the commonly-used Q5 Hi-Fidelity polymerase is around 1 in 10,000 bp (314). Computational and experimental methods of dealing with this error will be discussed in this section.

A common method of eliminating PCR error in diverse sequence sets (e.g bacterial metagenomics studies) is full-length clustering using an algorithm like USEARCH, which sorts all sequences by their copy number, using the most abundant sequences as centroids to agglomerate less-abundant sequences within a certain number of mutations (315). USEARCH has been incorporated into the SONAR pipeline (316), and into our custom pipeline. A conceptually similar homology-based centroid clustering approach, called AbCorrect, has also been developed (317). More advanced clustering schemes incorporating multiple alignment to a VDJ reference and incorporating quality scores also exist (318)(319).

Unfortunately, many clustering schemes cannot distinguish between a PCR-induced error and a rare somatically mutated variant within the clustering radius. A

clever graph-theoretic approach developed as part of the IgRepertoireConstructor package attempts to address this problem (320). First, each unique sequence is assigned to a vertex, and all vertices within a certain mutational distance are connected by an edge. Putative “real” sequences correspond to well-connected components called corrupted cliques within the graph; unfortunately, finding corrupted cliques is a computationally difficult (NP-hard (321)) problem. Within each clique, edges connecting sequences that differ at an SHM hotspot are treated distinctly from other mutations, in order to separate out “real” antibodies within a larger cluster from PCR errors. This algorithm, while clever, is also rather slow, though improvements published in 2016 are touted to decrease runtime from days to minutes for a typical dataset (322).

A widely-adopted experimental method to reduce error and the effects of amplification bias in transcriptomics is the incorporation of unique molecular identifiers (UMIs), short stretches of degenerate nucleotides, to cDNA during reverse transcription (323–325). Because the UMI is incorporated before subsequent amplification, all mutated “descendants” of that molecule erroneously created by PCR and sequencing will inherit the same UMI. Thus, clustering all reads sharing a UMI can not only be used to eliminate mutations, it can also be used to smooth over PCR-induced amplification biases that prevent accurate quantitation. Several recently-developed experimental pipelines, including MAF (326) and MIGEC (327)(328), incorporate UMIs into antibody repertoire sequencing to achieve significant decreases in error and accurate quantitation.



However, UMIs are not a panacea, and require appropriate experimental design for proper use. First, it is important to select a UMI scheme with sufficient diversity to uniquely tag every single mRNA molecule in a given sample (329)(330)(331). Insufficient UMI diversity will cluster together completely unrelated sequences, resulting in the elimination of the less-abundant one in the cluster. An easy method to increase UMI diversity is simply lengthening the UMI; however, this also makes the UMI more prone to mutations, which artificially increase the diversity of the sample (47)(332)(333). Methods of correcting mutations within the UMI include using histograms to find nearest neighbor distances between UMI variants and treating rarer variants as errors (334), and network-based methods (335). A final issue, which is especially noticeable when searching for rare variants, is the fact that the initial UMI tagging during reverse transcription may lack perfect efficiency; in some extreme cases, only 0.2% of the initial templates were successfully tagged (332)(333). No well-validated method for overcoming this bias has been implemented in antibody repertoire sequencing, though a recently-developed scheme in which single nucleic acid templates are encased in picoliter droplets and individually barcoded could theoretically eliminate competition between templates (336).

The use of UMIs has been widely adopted in repertoire sequencing, though some issues remain. Improved quantification and error correction comes at the potential of significantly decreased sequencing depth. Experiments in chapter 3 eschew the use of UMIs, partly because of our focus on deeply sampling the rare IgG4 subclass (necessitating high inputs of RNA), reliably separating the IgG4 subclass from the others

(necessitating an additional amplification step), and our desire to use phage display libraries for antigen-specific selection. For fine lineage analysis, we instead used USEARCH and a minimum read cutoff to filter out likely erroneous sequences, using a monoclonal amplicon subject to the same amplification steps to validate our method.

### **1.3.4 VDJ Assignment and Lineage Analysis**

Because the IgH variable locus is highly polymorphic and contains several similar families of sequences, the assignment of germline VDJ genes to a transcript is not trivial, even before taking SHM and the existence of N and P nucleotides into account. The original zebrafish repertoire sequencing experiment had a 99.6% V-gene assignment rate (293); with a more complex, mutated repertoire, this is bound to be worse. This is made even more difficult because several V gene alleles remain undiscovered (337). The simplest VDJ assignment algorithms are the alignment-based IgBlast (338) and IMGT-HighVquest (339), which are well-established tools with high throughput and simple web interfaces. Some alignment algorithms are particularly focused on the CDR3 region and purport to improve D gene assignment; examples include Ab-origin (340), which has a tweaked scoring scheme, and Joinsolver (341,342), which used Monte Carlo simulations to do the same. More advanced algorithms based on Hidden Markov Models (HMMs) include the SODA2 (343) and iHMMuneAlign (344), both of which use a negative binomial distribution for transition states, and the more complex multi-HMM based partis (345), which uses empirical distributions of VDJ recombination to better fit the data.

VDJ assignment is particularly important for analyzing lineages of related antibodies arising from a single naïve B cell clone (sometimes called a “clonotype”). Many pipelines (e.g. Change-O (116) and ClonalRelate (346)) assign sequences to lineages by first partitioning by assigned V gene, J gene, and CDR3 length, then separating out related sequences using a CDR3 distance cutoff between members of the lineage based on statistics. Alternately, a shared somatic mutation cutoff along the entire variable region can be used instead of CDR3 distance after partitioning, for example only allowing variable regions that share five mutations from their calculated germline sequence to share a lineage (296)(347) . Either approach would naturally separate related sequences mis-assigned to different V or J genes; this is of particular concern in sets of highly mutated sequences, like HIV bnAbs. One method to overcome this is to use phylogenetics on a database of V and J genes to improve the initial assignment, such as in IgSCUEAL (348). An alternative method is to assign sequences to lineages without first assigning V and J genes. This can be done, for example, by directly calculating a pairwise distance matrix between sequences (349), though this can become computationally intensive for large number of sequences.

In chapter 3, novel approaches are employed to deal with both VDJ assignment errors and clonal assignment. After performing germline and CDR3 assignments using IMGT-HighVquest, all sequences in the dataset are binned by V gene, J gene, and CDR3 length. However, in order to assign reads to lineages with more confidence, we chose to calculate the nearest-neighbor CDR3 nucleotide identity *between different* patients, rather than *within the same* patient, to set a per-patient cutoff. Because sequences

between patients should not be related, this can be used to bound the probability that any CDR3 similarity cutoff would erroneously cluster two unrelated sequences into a single lineage. The cutoffs for 99% confidence under this framework varied between 80.6% and 86.1%, quite similar to previously used values based on within-patient analysis. Secondly, to mitigate the effects of V- and J-gene mis-assignment and sequencing error, all clusters were subsequently sorted into superclusters solely based on CDR3 similarity, using the same per-patient cutoffs. Within each supercluster, clusters were ranked by read count, then only the top-ranked clusters within each supercluster were used for subclass connectivity analysis and lineage analysis through phylogenetic trees. This eliminated only ~10% of the total reads in each sample, but also eliminated anywhere from one-half to five-sixths of all clusters, significantly ameliorating the effects of sequencing error and germline mis-assignment on the construction of lineages.

Once sequences have been sorted into lineages, the next task is to draw a lineage tree using those sequences, which can offer insight into the temporal structure and evolution of the antibody response. Many tools from evolutionary biology can be adapted for this purpose, including maximum parsimony and maximum likelihood algorithms (350). Briefly, a maximum parsimony algorithm considers a set of sequences, and arranges them in a tree to minimize the total number of mutations (351), while a maximum likelihood method takes different site-specific mutation rates into account when drawing the tree, creating the most likely tree based on a calculated prior probability distribution (352). Both of these methods have been applied to antibody

sequences (116)(284)(353). Work in this field is ongoing, with more research needed to validate these methods and compare them against each other in immunological sequence datasets (47).

In order to ensure the robustness of our conclusions, we used both maximum parsimony (through ChangeO and PHYLIP's dnapars program) and maximum likelihood (through PASTA) algorithms to analyze anti-Dsg lineages containing both IgG1 and IgG4 sequences. Both methods generated trees with similar topologies, suggesting our conclusions were robust to the phylogenetic method selected. Additionally, ChangeO's particular implementation of maximum parsimony also calculates the sequence of inferred ancestors in the tree. This was particularly useful for determining the sequence of the most recent common ancestor (MRCA) in each lineage; these could then be synthesized (along with several other sequences in each tree) and tested by ELISA in order to trace the onset and development of anti-Dsg reactivity.

### **1.3.5 Conclusion**

It is clear that high-throughput sequencing can offer many insights into the antibody repertoire. In particular, the ability to infer lineage relationships within the antibody repertoire is intriguing, and is particularly relevant to the task of tracing the lineage of anti-Dsg B cells belonging to different antibody subclasses. By sequencing subclass-specific amplicons and sorting them into cross-subclass lineages, one can map the global relationship between IgA1, IgA2, IgG1, and IgG4. Furthermore, by combining this data with phage display as a means to isolate Dsg-specific heavy chains, we have the ability

to not only determine which sequences within the repertoire are autoreactive, but how they arose and how they relate to each other across antibody subclasses. The results of this repertoire sequencing study can be found in chapter 3.

## **1.4 Experimental Approach**

### **1.4.1 Rationale**

The focus of this thesis is to analyze the role of antibody subclass in the pathogenesis of PV. Because previous studies have shown that variable regions alone, either in Fab or single-chain variable fragment (scFv) form, are sufficient to cause blistering in model systems, there are broadly speaking two ways in which the subclass of an anti-Dsg antibody could be relevant to its role in pathogenesis. First, the subclass could be directly involved in the binding activity of the antibody, by altering the binding properties of the variable region. For example, a given anti-Dsg variable region could have different binding affinity when attached to either IgG1 or IgG4, due solely to structural differences in the constant region. The influence of the constant region on antibody affinity and pathogenicity is the subject of experiments found in chapter 2.

Alternatively, the subclass of an antibody could simply reflect the developmental history of the B cell synthesizing it. In other words, the structure or effector function of the heavy chain constant region that determines subclass may itself not be particularly important, but mapping the pathway leading autoreactive antibodies to occupy a particular antibody subclass may shed light on the processes involved in PV

pathogenesis. We can infer these pathways by analyzing clonal overlaps between anti-Dsg variable regions of different subclasses, tracing the lineages of autoreactive B cells as they switch from one subclass to another. In this way, analyzing subclass-specific B cell repertoires in PV will shed light on the mechanisms leading to the development of anti-Dsg antibodies across subclasses, and elucidate the contribution of class-switching to the pathogenesis of PV and antibody-mediated autoimmunity in general.

Experiments tracing subclass-specific B cell lineages in a panel of PV patients are found in chapter 3.

### **1.4.2 Chapter 2**

Experiments conducted in chapter 2 of this thesis analyze the possibility that epitope characteristics and of anti-Dsg antibodies are affected by subclass. While the variable region of an antibody is thought to be responsible for most of its binding properties, some studies have shown a role for the constant region in modulating reactivity (354). The anti-HIV bNAb 2F5 shows increased affinity and the ability to bind a new epitope when expressed as IgA2 than as IgG1 (355). Another group found that the catalytic DNA cleavage activity (356) and renal pathogenicity (357)(358) of the murine anti-nuclear antibody PL9-11 is influenced by its isotype. These findings raise the possibility that anti-Dsg3 IgG1 and IgG4 antibodies are closely related, and that class-switching from IgG1 to IgG4 is all that's needed to cause the same variable region to become pathogenic.

In chapter 2, we clone three anti-Dsg antibody variable regions previously isolated from PV patients onto IgG1 and IgG4 backbones. Then, we compare them pairwise for epitope preferences, ELISA and IIF titer, blistering ability on skin explants, and ability to internalize Dsg on the surface of keratinocytes. By cloning the same variable regions onto two different constant regions, we show that an antibody's affinity, pathogenicity, and ability to trigger Dsg internalization of an antibody is not significantly modulated by whether it expresses an IgG1 or IgG4 constant region.

### **1.4.3 Chapter 3**

Experiments conducted in chapter 3 trace anti-Dsg B cells across the IgG1, IgG4, IgA1, and IgA2 subclasses in the peripheral B cell repertoire of a panel of four PV patients. These experiments are conducted using RNA from peripheral blood mononuclear cells (PBMC) isolated from 50 mL of peripheral blood from each patient. This quantity of blood represents approximately 1% of the peripheral repertoire (given that humans harbor approximately 5L of blood) and 0.02% of the total B cell repertoire (given that 2% of the B cell repertoire is in the periphery (46)). The RNA is used to generate subclass-specific amplicons using a 3' primer binding to the hinge of the constant region and 5' primers binding to the beginning of FWR1 on the variable region. These amplicons are then used to create sequencing libraries for the Illumina MiSeq 2x300 bp paired-end sequencing protocol, and to create subclass-specific phage display libraries, which are subsequently panned against both Dsg1 and Dsg3 in order to isolate Dsg-reactive antibodies from each patient-subclass pair. Anti-Dsg clones isolated by



panning were subject to both ELISA and indirect immunofluorescence to confirm binding to desmoglein and recapitulation of the PV cell-surface binding phenotype, respectively.

Antibody phage display is a method of isolating antigen-specific antibodies by cloning heavy and light chains into a phagemid vector, which allows the antibody to be expressed as either a scFv or Fab on the surface of filamentous bacteriophage. A library of phage can then be subject to several rounds of selection and amplification against an immobilized antigen, isolating clones reactive to the antigen of interest (359). The technique is particularly useful to this study because it has the power to screen up to  $10^9$  different transformants in a single library, giving much higher throughput than single B cell cloning or heterohybridoma methods and conferring the power to find and characterize both rare and common clones. Phage display has been used to clone autoantibodies in acquired thrombotic thrombocytopenia purpura (360), idiopathic thrombocytopenia purpura (361), pemphigus foliaceus (362), and systemic lupus erythematosus (363)(364).

Previous phage display experiments in PV, in which all IgG subclasses were cloned into single libraries (making IgG1 and IgG4 indistinguishable), have been used to show that VH gene usage in anti-Dsg3 antibodies is highly restricted (208,365–367). Another set of phage-display derived pathogenic antibodies demonstrated a D/E-X-X-X-W motif in the heavy chain CDR3, suggesting that shared heavy chain sequence characteristics may be used to distinguish and target autoantibodies (209,368). Because of this restriction, and because of the fact that heavy chains are more diverse, are more

important for antigen binding, and generate traceable lineages, most of our analyses focus on heavy chain repertoires in PV.

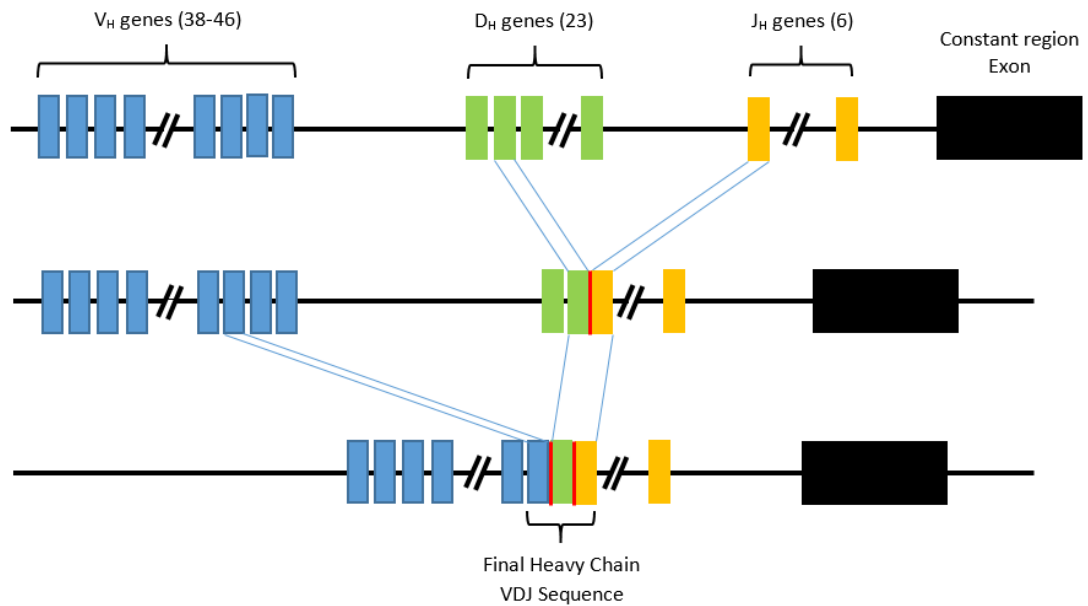
However, it is worth noting a key drawback of phage display is the random pairing of heavy and light chains in the library, since each is amplified separately from total RNA and PCR-overlapped into a single insert. Because of this, information about the actual heavy-light pairing is lost during cloning, creating the possibility that either a reactive heavy chain does not find a suitable light chain to pair with, or that a non-reactive heavy chain pairs with a light chain that makes it reactive. There are a few reasons to suggest that this is less of a problem than it initially appears. As would be suggested from its diversity, the heavy chain CDR3 has been experimentally demonstrated to be the main determinant of antibody specificity, with the light chain being less important. In one study, surface plasmon resonance assays showed that the heavy chain of an anti-thyroglobulin antibody recognizes its epitope in the complete absence of a light chain, showing the primacy of heavy chain interactions for binding (369). The primacy of the heavy chain was also shown by another group creating phage display libraries from a pool of several heavy chains with synthetic CDR3 regions and a single invariant light chain, and demonstrating that they harbor clones reactive against several haptens and protein antigens (370)(371). Another group showed that transgenic mice expressing a single heavy chain V gene, but the full complement of D and J genes, developed high-affinity antibodies in response to several different injected antigens, despite restricted combinatorial diversity, showing that the combinatorial and junctional diversity within the heavy chain CDR3 was sufficient to generate such responses (372).

By tracing the B cell repertoire across multiple subclasses, we were able to determine that the IgG4 anti-Dsg lineages tended to be solitary, containing only IgG4 clones. Furthermore, when IgG1 and IgG4 shared a lineage, they often diverged from a common ancestor early, rather than showing a direct switch between IgG1 and IgG4. This is in contrast to IgA1 and IgA2, which commonly showed IgA1 to IgA2 direct-switch relationships. We also found that IgA subclasses tended to switch directly from IgG subclasses, rather than the other way around, suggesting that IgA may serve as a disease-modifying factor in PV, rather than serving as a source of anti-Dsg IgG clones. These results and analysis can be found in chapter 3.

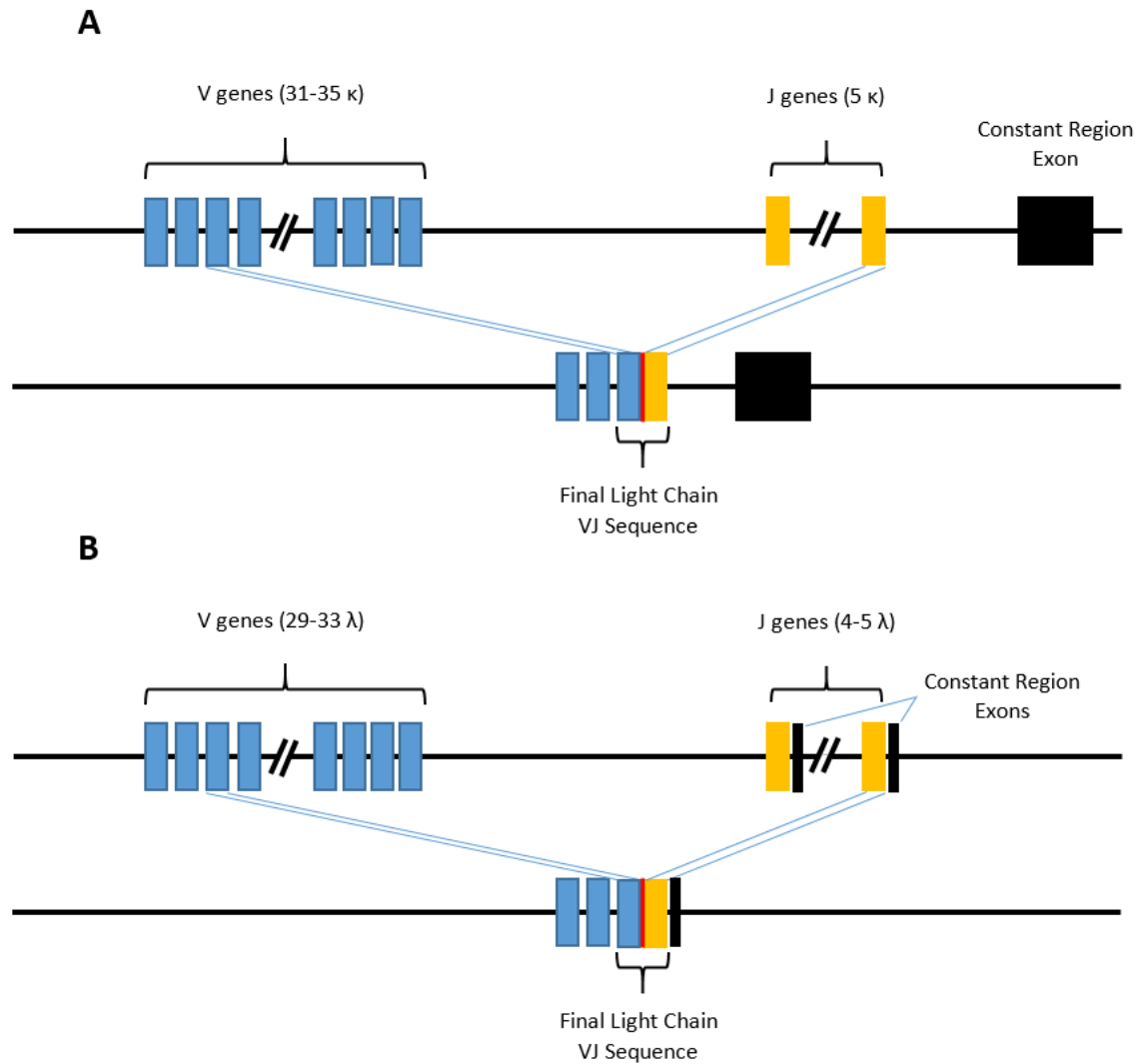
## 1.5 Figures and Tables

Disease	Subtype	Autoantigen	Histology	Clinical Appearance
Pemphigus Vulgaris (PV)	Mucosal	Dsg3	-Suprabasal acantholysis -Igg deposition on cell surface by DIF	-Erosions and fragile blisters restricted to mucosal surfaces, particularly the oral mucosa
	Mucocutaneous	Dsg3 and Dsg1	-Eosinophilic abscesses are common -Suprabasal acantholysis, but more subtle than PV	-Erosions and blisters found in both skin and mucous membranes -Rare variant of PV (1-2%)
Pemphigus Vegetans	Sporadic	Dsg3	-Erosions that later organize into fungoid vegetations	-Commonly found on flexural surfaces
		Dsg1	-Superficial clefting	-Scaly, crusted erosions from superficial vesicles
Pemphigus Foliaceus (PF)	Pemphigus Erythematosus	Dsg1 (also ANA)	-In addition to normal features of PF, IgG and C3 in basement membrane and intercellular	-Considered a localized variant of PF -Scaly lesions on face can resemble malar rash of SLE
	Endemic (Brazilian)/ Fogo Selvagem (373-375)	Dsg1	-Similar to sporadic PF	-Similar to sporadic PF
	Endemic (Tunisian) (201,376)	Dsg1	-Similar to sporadic PF	-Similar to sporadic PF
	Endemic (Colombian)(377-379) (380)	Dsg1 +/- desmoplakin 1 +/- periplakin +/- envoplakin +/- neural autoantigens	-Igg deposition on keratinocyte skin surfaces -Igg, Iga, Igm and C3 on basement membrane zone (40%) -Subcorneal and intraepidermal acantholysis	-Hyperpigmented plaques/macules on face and trunk -Butterfly-like macular rash with scaling and erosions on face -Extensor weakness caused by neural autoantibodies
Iga Pemphigus	Subcorneal Pustular Dermatitis	Desmocollin 1	-Iga deposition in superficial layers of skin -Subcorneal neutrophilic pustules	-Flaccid vesicles or pustules, in an annular or circinate pattern
	Intraepidermal Neutrophilic Iga Dermatitis	Unknown (some cases Dsg3/ Dsg1)	-neutrophilic pustules in stratum spinosum -Iga deposition throughout entire epidermis	-Flaccid vesicles or pustules, in an annular or circinate pattern -Sunflower-like arrangement of pustules
Paraneoplastic Pemphigus (PNP)		Dsg3, Dsg1, Plectin, Desmoplakin 1, Desmoplakin 2, BP230, Envoplakin, Periplakin	-Highly variable, due to wide range of autoantigens -Often show suprabasilar acantholysis, along with individual keratinocyte necrosis and lymphocytic infiltrate	-Severe stomatitis all over oropharynx, and other mucosal surfaces -Cutaneous findings can include flaccid blisters, tense blisters, lichenoid eruptions, erythema multiforme-like lesions depending on case

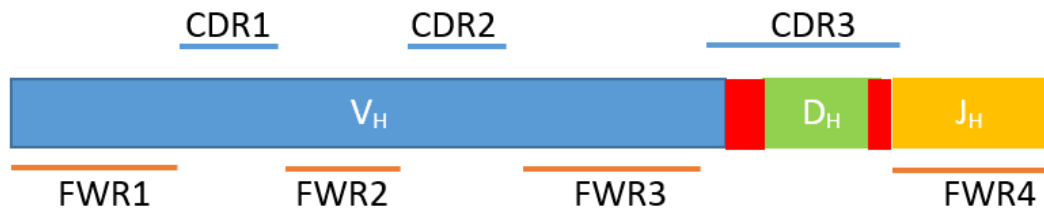
Table 1. Summary of Pemphigus and its Subtypes. Information from (381), except where indicated.



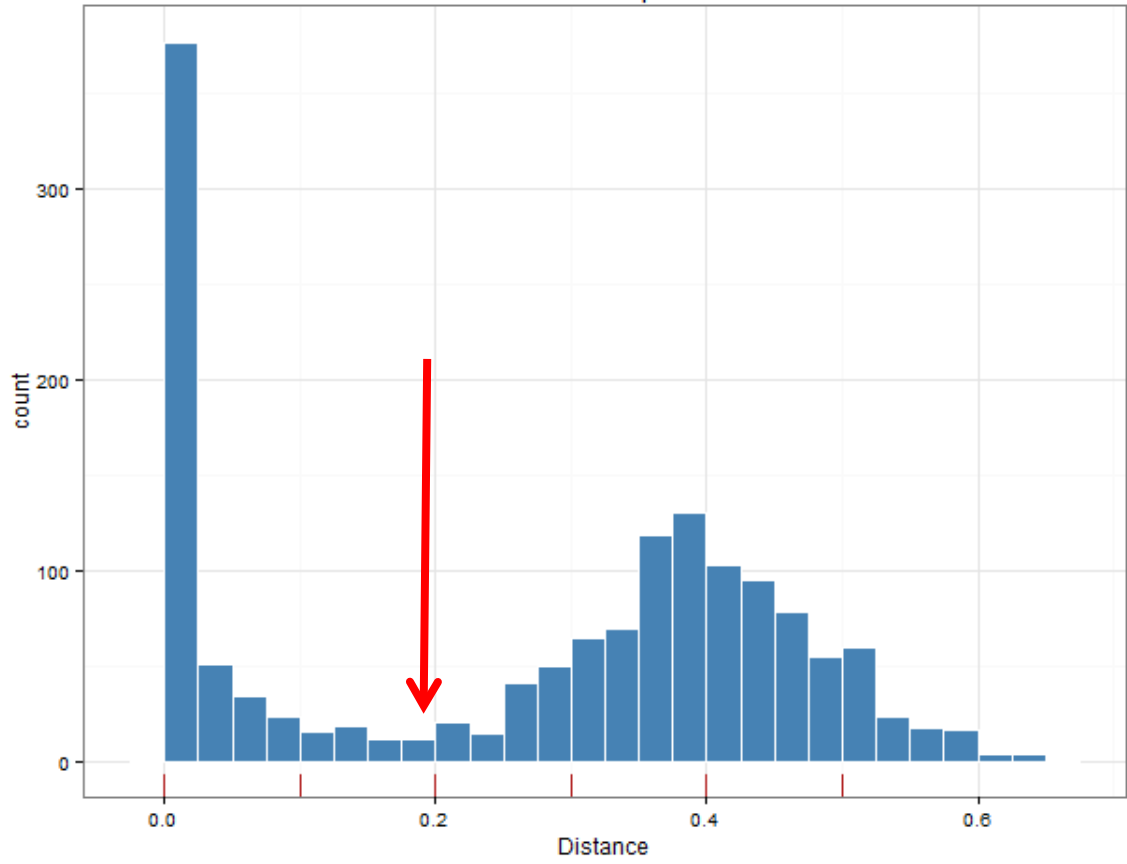
**Figure 1-1. Diagram of VDJ Recombination at the Human Heavy Chain Locus.** A D and J segment are randomly selected then joined together by RAG1/2, with nucleotides at the junction (in red) added by TdT or removed by exonuclease chewback. A similar process then occurs with a randomly selected V and the already-joined DJ segment, creating the final VDJ variable chain.



**Figure 1-2. Diagram of VJ Recombination at the Human (A) Kappa and (B) Lambda Locus.** A V and J segment are randomly selected then joined together by RAG1/2, with nucleotides at the junction (in red) added by TdT or removed by exonuclease chewback. In the Kappa locus, there's a single constant region exon on the 3' end of the locus, while in the Lambda locus, there's a constant region exon immediately 3' of each J gene.

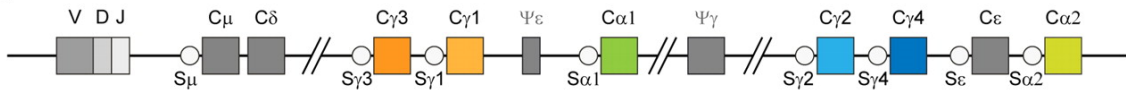


**Figure 1-3. CDRs and FWRs are Interspersed among the V, D and J Genes of the Variable Heavy Chain.** In particular, the V gene contributes to both the CDR1 and CDR2, but all three genes contribute to the CDR3. Red indicates junctional nucleotides added during VDJ recombination (see Figure 1-1).



**Figure 1-4. Clonality Plot.** 2000 sequences obtained by deep sequencing from PV8 were randomly selected. After binning sequencing by V, J and CDR3 length, the nearest-neighbor hamming distance between CDR3s are measured, in order to separate clonally related sequences from unrelated sequences. The red arrow indicates the minimum in the histogram, which would be used to set a CDR3 similarity cutoff for clonality. Distance = (Hamming distance, or # of substitutions)/(CDR3 Length)





**Figure 1-5. Diagram of Human IgH Constant Region Locus.** Each exon encodes a different heavy chain constant region, as indicated. Circles immediately adjacent to each exon represents switch regions, which serve as substrates for AID to initiate class switching. Note that class switching occurs from left to right, so for example, an IgM to IgG4 switch is possible, but not vice versa, as the intervening DNA is excised. Note that as displayed, there is no switch region in front of IgD, consistent with its co-expression with IgM. However, a small proportion of B cells do switch to expressing solely IgD, made possible due to the presence of a rudimentary switch region between IgM and IgD (121,373).

Adapted from: Figure 4a of

Berkowska MA, Driessen GJA, Bikos V, Grosserichter-Wagener C, Stamatopoulos K, Cerutti A, et al. Human memory B cells originate from three distinct germinal center-dependent and -independent maturation pathways. *Blood*. 2011 Aug 25;118(8):2150–8.

## **CHAPTER 2:**

### **Swapping the Constant Region in Pathogenic IgG1 and IgG4 Autoantibodies in Pemphigus Vulgaris Does Not Alter Blister-Inducing Ability or Epitope Preference**

#### **2.1 Abstract**

Pemphigus vulgaris (PV) is characterized by IgG1 and IgG4 autoantibodies to desmoglein 3 (Dsg3), causing suprabasal blistering of skin and mucous membranes. Both IgG1 and IgG4 appear during active disease, whereas IgG1 can be associated with remittent disease; this partially reciprocal relationship, and the relative positions of heavy chain constant region loci, suggests the possibility that IgG1 clones switch to IgG4 clones over the course of disease. To investigate the possible role of the IgG constant region on modulating blister-inducing ability and binding properties in anti-Dsg3 autoantibodies, we tested whether three pathogenic PV mAbs from three different patients demonstrate differences in antigen affinity, epitope specificity, or blister-inducing ability in model systems when expressed as IgG1 or IgG4. F706 anti-Dsg3 IgG4 and F779 anti-Dsg3 IgG1, previously isolated as heterohybridomas, and Px43, a monovalent cross-reactive anti-Dsg3/Dsg1 IgG antibody isolated by phage display, were subcloned to obtain paired sets of IgG1 and IgG4 mAbs. Using ELISA and cell surface staining assays, F706 and F779 demonstrated similar antigen binding affinities of IgG1 and IgG4. Px43 showed 3 to 8-fold higher affinity of IgG4 versus IgG1 by ELISA, but identical binding affinities to human skin, perhaps due to targeting of a quaternary epitope best displayed in tissues. All 3 monoclonal antibody (mAb) pairs targeted the same extracellular cadherin domain on Dsg3, caused Dsg3 internalization in primary human keratinocytes, and caused suprabasal blisters in human skin at comparable doses. This work suggests that switching the IgG1 and IgG4 subclasses of pathogenic PV mAbs does not directly affect their antigen binding or pathogenic properties.

## 2.2 Introduction

Pemphigus vulgaris is a potentially life threatening autoimmune blistering disease caused by IgG autoantibodies against Dsg3 +/- Dsg1, desmosomal adhesion proteins of keratinocytes (374). In mucosal PV, autoantibodies react against Dsg3, which is the major Dsg isoform expressed in basal keratinocytes of the mucosa, leading to suprabasal blistering. In mucocutaneous PV, autoantibodies target both Dsg3 and Dsg1, causing suprabasal blisters in both mucosa and skin. Antibody cloning studies from PV patients have shown that a single mAb can bind Dsg3 only, or both Dsg3 and Dsg1 (375)(208)(376). Epitope mapping studies indicate that autoantibodies from pemphigus patients with active disease bind near the N-terminus of Dsg3, where residues important for cis- and trans-adhesive interactions reside (225,376).

Although monovalent anti-Dsg antibody variable region fragments are sufficient to cause skin blisters in the absence of the constant region (377), preferential use of IgG subclasses has been identified in pemphigus patients. In active disease, anti-Dsg3 IgG4 autoantibodies are predominantly found in PV sera, followed by IgG1, and occasionally IgG2 and IgG3 (243). In contrast, patients in remission and sometimes healthy relatives of PV patients can demonstrate anti-Dsg3 IgG1 (245,378). These early studies in the field suggested that IgG4 antibodies could induce blisters but IgG1 antibodies could not. Subsequently, heterohybridomas produced from PV patients have shown that pathogenic anti-Dsg3 mAb can derive from either the IgG1 or IgG4 subclass (367,375,376).

While it is clear that IgG4 is associated with active disease in PV, it is unknown whether the same variable region of a pathogenic IgG4 mAb, expressed on an IgG1 constant region, directly affects antibody function, through altering antibody affinity, epitope binding characteristics, Dsg3 internalization, or blister-inducing ability. In the study presented here, we selected three representative pathogenic PV mAbs cloned from 3 different PV patients: F706, an

anti-Dsg3 IgG4 isolated by heterohybridoma, F779, an anti-Dsg3 IgG1 isolated by heterohybridoma, and Px43, an anti-Dsg3/Dsg1 monovalent IgG antibody of unknown subclass isolated by phage display, and generated pairs of recombinant IgG1 and IgG4 antibodies expressing identical variable regions. To determine whether the constant region modulates affinity, epitope specificity, or pathogenic properties of anti-Dsg variable regions, we compare these antibodies pairwise by ELISA titer, IIF titer, pathogenicity in skin explants, and Dsg internalization ability.

## **2.3 Materials and Methods**

### **2.3.1 Human Anti-Dsg3 mAb Cloning**

The three patients from whom the antibodies were cloned had active, extensive PV involving both mucosa and skin. The diagnosis of PV was confirmed by histology and immunofluorescence. All patients were untreated at the time of blood draw. F706 IgG4 and F779 IgG1 were isolated from two different patients by heterohybridoma as described (367,379). Px43 was isolated from a third PV patient by phage display (208).

### **2.3.2 Recombinant mAb Subcloning, Expression, and Purification**

Primers were designed to amplify the F706 and F779 heavy chain (HC) and light chain (LC) variable domains for subcloning into the pHc or pLc vector (380,381). Primers are designed to bind to the 5' UTR (leader) and the 3' end of the JH-gene, while light chain primers bind to the 5' UTR (leader) and the constant region. Primer sequences were as follows (restriction sites underlined):

5' F706HC NheI: TCTAGCTAGCCGCCACCATGGACTGGACCTGGAGGGTCTT

3' F706HC HindIII: TAGGGCAAAGCTTGCTGAGGAGACGGTGACCAGG

5' F706LC NheI: TCTAGCTAGCCGCCACCATGAGGCTCCTTGCTCAGCT

3' F706LC NotI: TAGGGCGCGGCCGCAGTTCGTTTGATTCCACCTT

5' F779HC NheI: TCTAGCTAGCCGCCACCATGGACTGGACCTGGAGGGTCTT

3' F779HC HindIII: TAGGGCAAAGCTTGCTGAGGAGACGGTGACCAGG

5' F779LC NheI: TCTAGCTAGCCGCCACCATGAGGCTCCTTGCTCAGCT

3' F779LC NotI: TAGGGCGCGGCCGCAGTTCGTTTGATCTCCAGCTT

Recombinant DNA plasmids encoding cognate heavy and light chains were co-transfected into CHO-K1 cell lines using Lipofectamine 2000 (Invitrogen, Grand Island, NY), with IgG1 and IgG4 heavy chains cloned into separate lines. Transfectants were cloned by limiting dilution in selective media (RPMI containing 10% fetal bovine serum, 10µg/mL puromycin, 800µg/mL G418) to establish permanent cell lines. The antibody supernatant was collected and purified by affinity chromatography on a Sepharose protein G column with acid glycine elution. F71 IgG1 mAb (targeting an unknown antigen) and anti-HIV b12 IgG4 mAb were used as isotype antibody controls for some *in vitro* experiments.

Px43 IgG1 and IgG4 were cloned as previously described (368)(382). Briefly, heavy and light chains were individually amplified from Px43 scFv in the pComb3X vector, then cloned into the PIGG vector (383). The heavy chain was amplified to introduce a SacI cloning site on the 5' end and an Apal cloning site on the 3' end (the endogenous Apal site in the CH region of IgG). After PCR purification using the Qiaquick PCR purification kit (Qiagen, Valencia, CA), digestion using SacI and Apal (New England BioLabs, Ipswich, MA) and gel purification using the Qiaquick gel extraction kit (Qiagen), the heavy chain fragment was subcloned into the SacI-Apal site on the PIGG vector. The light chain was amplified from the same template, using a 5' primer that

abolishes the SacI site and introduces a HindIII site, and a 3' primer that binds to the 3' end of the J gene. The light chain constant region was amplified from a separate Fab-expressing vector, then the light chain variable and constant regions were joined by overlap PCR. The overlap PCR reaction was PCR purified, digested with HindIII and XbaI (New England BioLabs), and gel purified (Qiagen) before being subcloned into the HindIII-XbaI site of the PIGG vector. A complete set of primers is available in (368).

Px43 IgG1 and IgG4 were expressed and purified as previously described(368). Briefly, endotoxin-free plasmid maxipreps (Qiagen) were transfected into five 10cm culture dishes of HEK293T cells at approximately 70% cell density and DMEM plus 10% ultralow IgG fetal bovine serum (Gibco/Life Technologies, division of ThermoFisher Scientific, Waltham, MA) using jetPEI reagent (Polyplus-Transfection SA, Illkirch, France) according to manufacturers instructions. Antibody was harvested on Day 3 and Day 6 from culture supernatant, by centrifuging non-adherent cells then neutralization of media by 1/100 volume 1M Tris pH 7.5. Antibody was purified using Protein A agarose (ThermoFisher Scientific). After 2 hours of rotation at room temperature, agarose was transferred to disposable chromatography columns, washed with 20 column volumes of PBS, eluted with 6 column volumes of 100 mM glycine pH 3, then neutralized with 1/10 volume Tris pH 7.5. Concentration was determined by non-reducing SDS-PAGE followed by Coomassie Blue staining, using an IgG standard.

### **2.3.3 Epitope Mapping**

All epitope mapping vectors were generously provided by Masayuki Amagai. For epitope mapping, extracellular full length and domain-swapped Dsg1/Dsg3 or Dsg2/Dsg3 molecules were produced in the baculovirus-Sf9 system as previously described (226,384). For each experiment, culture supernatants containing domain-swapped constructs were incubated with

antibody followed by precipitation using protein A or protein G agarose. After boiling in Laemmli sample buffer, beads are run on SDS-PAGE gel, then transferred to polyvinylidene difluoride (PVDF) membrane. Immunoblotting was conducted using alkaline phosphatase-conjugated goat antihuman IgG (Zymed Laboratories, San Francisco, CA) for the Dsg1/Dsg3 swapped molecules (because their C-terminus consists of the IgG1 constant region) or anti-E HRP (ab19400, Abcam, Cambridge, MA) for the Dsg2/Dsg3 swapped molecules (because they express a C-terminal E-tag)

### **2.3.4 Antibody ELISA**

Dsg1 or Dsg3 ELISA (MBL International, Woburn, MA or Euroimmun, Luebeck, Germany) was used to measure relative binding affinities of IgG subclasses. Serial dilutions of IgG1 and IgG4 antibodies were incubated with Dsg1 or Dsg3 coated plates for 1 hour at room temperature using kit sample buffer or TBS buffer containing 1mM CaCl<sub>2</sub> and 5% dry milk. After washing, bound antibody was detected using HRP-conjugated mouse monoclonal anti-human IgG F(ab')<sub>2</sub> (MBL International, Woburn, MA) or mouse monoclonal anti-human lambda (clone JDC-12, Abcam). After 1 hour incubation at room temperature, plates were washed, developed with 3,3',5,5' tetramethylbenzidine substrate, and optical density determined at 450 or 490 nm using an automated plate reader.

Relative affinity was determined using regression analysis of double-reciprocal plots of optical density versus antibody concentration. The relative constant was calculated as  $(1/y \text{ intercept}) * \text{slope}$ .

### **2.3.5 Direct and Indirect Immunofluorescence**

Direct immunofluorescence (DIF) was performed on frozen human skin sections after mAb injection (see methods subsection on the human skin blistering assay – the same sample was split after injection for DIF and H&E). Bound antibody in human skin was detected with Alexa Fluor 594 anti-human IgG (Invitrogen, Grand Island, NY) in TBS buffer containing 1mM CaCl<sub>2</sub> and 1% BSA and visualized using an Olympus BX61 microscope.

Indirect immunofluorescence (IIF) was performed by incubating serial dilutions of mAb on frozen normal human skin sections (obtained through the Penn Skin Disease Research Center) in TBS buffer containing 1mM CaCl<sub>2</sub> and 1% BSA. Bound antibody was subsequently detected with FITC-conjugated anti-human lambda light chain (clone JDC-12, Abcam) and visualized using an Olympus BX61 microscope. The titer is reported as the last dilution at which epidermal cell surface staining is clearly positive.

### **2.3.6 Immunoreactivity Using Live Cell Staining ELISA**

Human oropharynx epithelial tumor cell line, UM-SCC-47 (SCC-47), was obtained from Dr. Thomas Carey, University of Michigan. Cell line was grown in 10% FBS-DMEM medium containing 10% heat inactivated FBS, 2mM L-glutamine, 0.1 mM sodium pyruvate, 0.1 mM non-essential amino acids, and 10 µg/ml gentamicin. A live cell based ELISA was employed using SCC-47 cells seeded in a 96-well flat bottom plate for two days. Serial dilutions of PV IgG1 and IgG4 mAbs and isotype control antibodies were added to the cells for one hour. After washing, bound antibody was detected using biotin-conjugated goat anti-kappa chain antibody (1:5000) (Southern Biotech, Birmingham, AL) for one hour. The plate was washed and HRP-conjugated streptavidin (1:10,000) (SouthernBiotech) was added and incubated for 45 minutes. After a final



wash, o-phenylenediamine (OPD) substrate (Sigma, St. Louis, MO) was added, and the plate was incubated for 20 min and read at 490nm.

### **2.3.7 Human Skin Blistering Assay**

All experiments using primary human skin samples were obtained using procedures approved by the Institutional Review Board. Fresh normal human skin specimens (otherwise discarded from excisional skin surgery) were obtained from the Penn Skin Disease Research Center. Skin was trimmed of fat and cut into 5 mm sections. 25-50 µg of purified isotype mAb, with or without 100 ng of *Staphylococcus aureus* exfoliative toxin A (ETA, Toxin Technology, Inc., Sarasota, FL) to inactivate Dsg1, was injected into skin sections, followed by incubation of skin pieces on transwell inserts (Corning, Lowell, MA), with defined keratinocyte serum free media (Invitrogen, Grand Island, NY) containing 1.2 mM CaCl<sub>2</sub> in the outer compartment. Skin was harvested at 16-24 hours for direct immunofluorescence (see methods subsection on immunofluorescence) by embedding in OCT compound (TissueTek, Sakura, Netherlands) and for histology by fixation in 10% phosphate buffered formalin.

Primary human epidermal keratinocytes (PHEK) obtained from the Penn Skin Disease Research Center were treated with 50 µg/ml anti-Dsg3 mAbs for the indicated amount of time during calcium-induced (0.07 mM shift to 1.2 mM calcium) desmosome assembly as described (221). At 8 hours, cells were double-labeled with Alexa Fluor 594 anti-human IgG and mouse anti-human Dsg3 (5G11) followed by Alexa Fluor 488 anti-mouse IgG, and nuclei were counter-stained with DAPI.

## **2.4 Results**

### **2.4.1 IgG1 and IgG4 Variants of Pathogenic PV mAbs Recognize the Same Antigenic Epitopes**

Because the hinge region of IgG1 differs from IgG4, creating differences in the orientation of the variable regions (253), Dsg3 affinity, epitope specificity, or blister-inducing ability could be affected by the presence of the IgG1 versus IgG4 Fc. To explore this further, IgG1 and IgG4 variants of pathogenic PV mAbs, bearing the identical variable region, were generated.

To determine whether the IgG1 or IgG4 constant region affects epitope specificity, we first tested F706 and F779 IgG1 and IgG4 mAbs against recombinant purified Dsg1 and Dsg3 antigens. ELISA showed that F706 and F779 IgG1 and IgG4 mAbs specifically bound to Dsg3 but not Dsg1 (Figure 2-1A), indicating that the switch of constant region from IgG1 to IgG4 or from IgG4 to IgG1 in these pathogenic antibodies did not affect Dsg isoform specificity. To further define the epitope specificities, we performed conformational epitope mapping of IgG1 and IgG4 mAbs using Dsg3 domain-swapped chimeric molecules. F706 and F779, both IgG1 and IgG4, mapped to the first 161 amino acids of Dsg3 (Figure 2-1B). Similarly, Px43 IgG1 and IgG4 mapped to the same epitope within the first 101 amino acids of Dsg3. Thus, class-switched IgG1 and IgG4 variants do not change the variable region epitope specificity.

### **2.4.2 PV IgG1 versus IgG4 mAbs Demonstrate Similar Antigen Affinities**

To determine whether IgG1 or IgG4 constant regions affect antigen affinity, we evaluated IgG1 and IgG4 variants of pathogenic PV mAbs by ELISA and keratinocyte cell surface staining. IgG1 and IgG4 variants of F706 and F779 show a similar level of binding to Dsg3 by

ELISA (Figure 2-2A). The apparent affinity of F706 IgG1 and IgG4 mAbs for Dsg3 is 0.10  $\mu$ M and 0.067  $\mu$ M, respectively, and 0.015  $\mu$ M or 0.021  $\mu$ M for F779 IgG1 or IgG4, respectively.

Consistent with these data, binding isotherms of F706 and F779 IgG1 and IgG4 mAbs to Dsg3+ human oropharynx keratinocyte tumor cells by immunolabeling are similar (Figure 2-2B). Px43 IgG1 demonstrates 8-fold and 3-fold lower affinity for Dsg3 and Dsg1 than Px43 IgG4 by ELISA (Figure 2-2A); however their human skin IIF titers are identical at 0.5 ng/ $\mu$ L (Figure 2-2C).

### **2.4.3 PV mAb IgG1 Versus IgG4 Demonstrate Comparable Blister Forming Ability in Human Skin Explants**

We next investigated the blister-inducing ability of IgG1 and IgG4 recombinant mAbs, to determine whether IgG subclass could play a role in modulating the pathogenic effects of Dsg3-reactive variable regions. In skin, Dsg1 can provide compensatory adhesion if Dsg3 is inactivated (235,385). Consequently, human skin assays to determine blister-inducing ability of anti-Dsg3 mAbs require inactivation of Dsg1, for example by co-injection of *Staphylococcal* exfoliative toxin A (ETA), which specifically cleaves Dsg1 (386). We injected 25-50  $\mu$ g of anti-Dsg3 IgG1 or IgG4 mAbs, plus or minus ETA as required, into human skin sections, followed by incubation for 16-24 hours prior to harvest for immunofluorescence and histology. F706, F779, and Px43 all caused suprabasal blisters at comparable doses of IgG1 and IgG4 (Figure 2-3A). A similar extent of blistering was observed at both 25  $\mu$ g and 50  $\mu$ g doses (Figure 2-3B). Direct immunofluorescence of injected skin sections showed comparable binding of PV IgG to the cell surface of keratinocytes (Figure 2-3C).

Previously, pathogenic but not non-pathogenic human PV mAbs were shown to cause Dsg3 internalization in primary human keratinocytes (221). Consistent with these studies, primary human keratinocytes treated with F706, F779, or Px43 IgG1 and IgG4 exhibited

depletion of cell surface Dsg3 within 4 hours after antibody treatment, concomitant with an increase in cytoplasmic vesicular staining of Dsg3 (Figure 2-4). Thus, IgG1 and IgG4 variants of F706, F779, and Px43 comparably induce loss of cell surface Dsg3 in primary human epidermal keratinocytes and suprabasal blisters in human skin.

## 2.5 Discussion

This study demonstrates that epitope specificity, antigen affinity, and blister-inducing ability of PV mAbs are not directly affected by the presence of the IgG1 versus IgG4 constant region. Although our conclusions are limited by the study of only three mAbs, these mAbs were cloned from 3 different patients and represent three distinct types of pathogenic PV mAbs that have been described in the literature, including anti-Dsg3 IgG1, anti-Dsg3 IgG4, and a cross-reactive anti-Dsg3/Dsg1 IgG.

All PV mAbs tested in this study bind the Dsg3 extracellular cadherin (EC) 1-2 domains that are crucial for cis- and trans-adhesive interactions of the Dsgs. This result agrees with a previous large scale mapping study of PV sera in which 91% of PV sera mapped to the Dsg3 EC1 domain (226).

Although F706, F779, and Px43 all demonstrated comparable antigen affinity of IgG1 and IgG4 variants by ELISA and/or IIF staining, Px43 IgG1 and IgG4 were notable for demonstrating differing results by ELISA and IIF. Px43 was previously shown to recognize only mature conformational Dsg3 (387) and by surface plasmon resonance binds Dsg3 with complex kinetics that best fit to a conformational change model (367). It is possible that Px43 binds a quaternary epitope in Dsg3 and Dsg1 that is predominant in native human skin but may be differentially displayed in ELISA, which may account for the difference in relative affinities observed between IgG1 and IgG4 variants. Nevertheless, these results indicate that the binding

affinity of Px43 variable regions to Dsg3 in human skin is not significantly altered by the presence of the IgG1 versus IgG4 constant region.

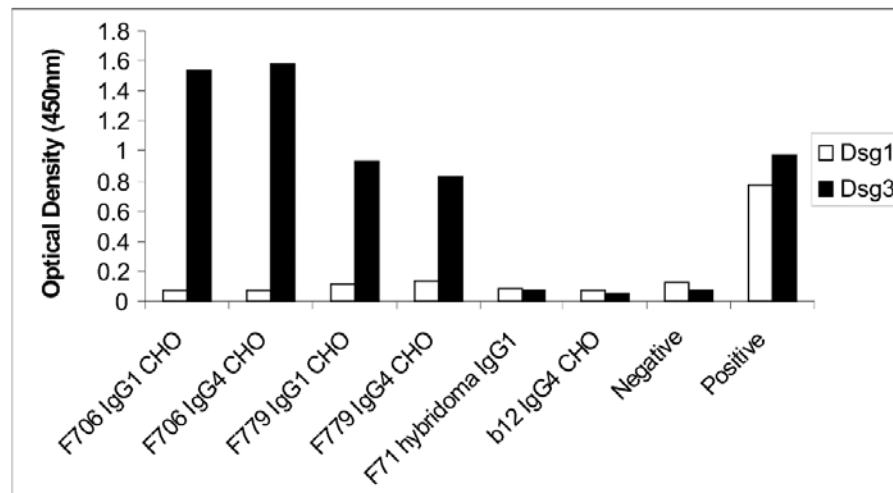
No change of blister-inducing ability among PV IgG1 and IgG4 mAb variants was observed, either in regard to Dsg3 internalization or the ability to cause suprabasal blisters in human skin. These data are consistent with the finding that the constant region is not required for blister formation in PV (377) or for Dsg3 internalization after monovalent PV mAb binding (221). Our data now further indicate that the IgG1 versus IgG4 constant region does not modulate the functional effects of pathogenic PV mAbs in regard to either Dsg3 internalization or suprabasal blister induction.

With no difference in pathogenicity, the question arises as to why IgG4 is preferentially associated with active disease. IgG4 is typically observed in conditions of chronic antigen stimulation, such as in individuals undergoing allergic desensitization or beekeepers (253). Class switch recombination is a complex process regulated by cytokines (388). IL-4 has a central role in stimulating antibody class switch (389). CD40 and IL-4 can initiate class switch by inducing enzymes and other transcripts required for site-specific DNA recombination. IgG1 is the most abundant serum IgG subclass, in part due to its proximal location within the constant region locus. Repeated antigenic exposure can encourage subsequent isotype switching. IL-4 and IL-13 promote isotype switching first to IgG4 and subsequently to IgE (390), whereas IL-10 and IL-21 potentiate IL-4-induced switching to IgG4 over IgE (391,392). Upregulation of the Th2 cytokines IL-4, IL-5, IL-10, and IL-13 has been described in pemphigus patients (393,394). These cytokines would promote an IgG4>IgG1 serum antibody profile, coinciding with the spectrum of observed autoantibody isotypes in patients with active disease. A deeper understanding of how pathogenic variable regions segregate among the IgG subclasses would require isotype-specific antibody cloning, as seen in chapter 3.

In summary, we have found that subclass-switching between IgG1 and IgG4 constant regions has no significant effect on epitope specificity, antigen binding affinity, or blister-inducing ability in 3 human PV mAbs isolated from 3 different PV patients. Although limited by the study of only three antibodies, this study provides the first direct evidence that the immunochemical and pathogenic properties of PV pathogenic variable regions are not significantly modulated by antibody subclass, supporting the conclusion that the antibody variable region is most important for determining pathogenicity in PV. The possible ontogeny of the variable region in PV, and specifically the lineage relationships between anti-Dsg variable regions in PV, will be further investigated in chapter 3 of this thesis.

## 2.6 Figures and Tables

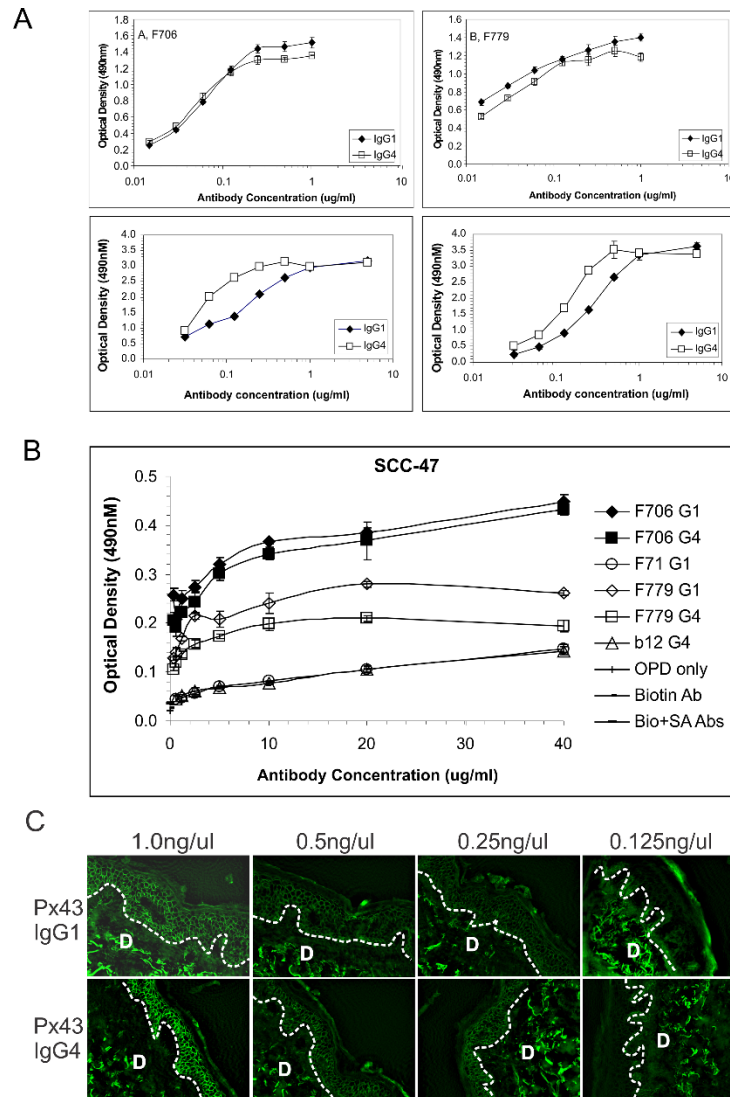
A



B

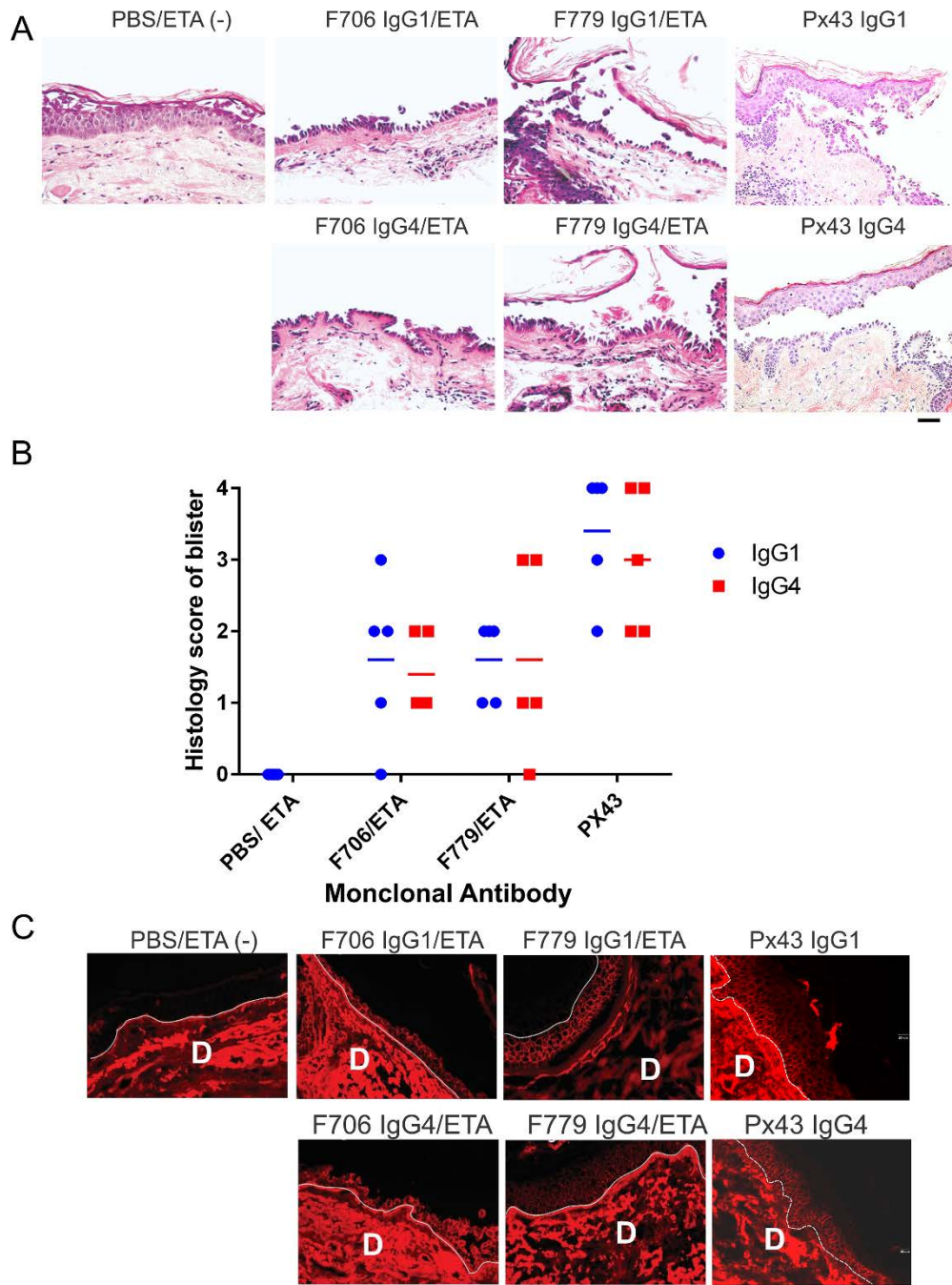
	F706 IgG1	F706 IgG4	F779 IgG1	F779 IgG4		Px43 IgG1	Px43 IgG4
hDsg1	0	0	0	0	hDsg1	+	+
hDsg3	+	+	+	+	hDsg3	+	+
hDsg3 (1-161)	+	+	+	+	hDsg3 (1-101)	+	+
hDsg3 (1-403)	+	+	+	+	hDsg3 (95-221)	0	0
hDsg3 (163-566)	0	0	0	0	hDsg3 (212-328)	0	0
hDsg3 (405-566)	0	0	0	0	hDsg3 (320-440)	0	0
Deduced epitope	1-161	1-161	1-161	1-161	Deduced epitope	1-101	1-101

**Figure 2-1. IgG1 and IgG4 PV mAbs target epitopes within the same Dsg3 extracellular domains.** (A) F706 and F779 mAbs bind desmoglein 3 (Dsg3) but not Dsg1. Subclass-switched anti-human Dsg3 F706 and F779 IgG1 and IgG4 were tested by ELISA for binding to Dsg3 and Dsg1. Negative and positive controls were provided by the manufacturer. F71 IgG1 and b12 IgG4 antibodies served as isotype controls. (B) Deduced epitopes of IgG1 and IgG4 variants of PV mAbs show that mAb pairs recognize epitopes within the same Dsg3 extracellular domains. Assay was considered positive if immunoblot showed a band at the correct size.

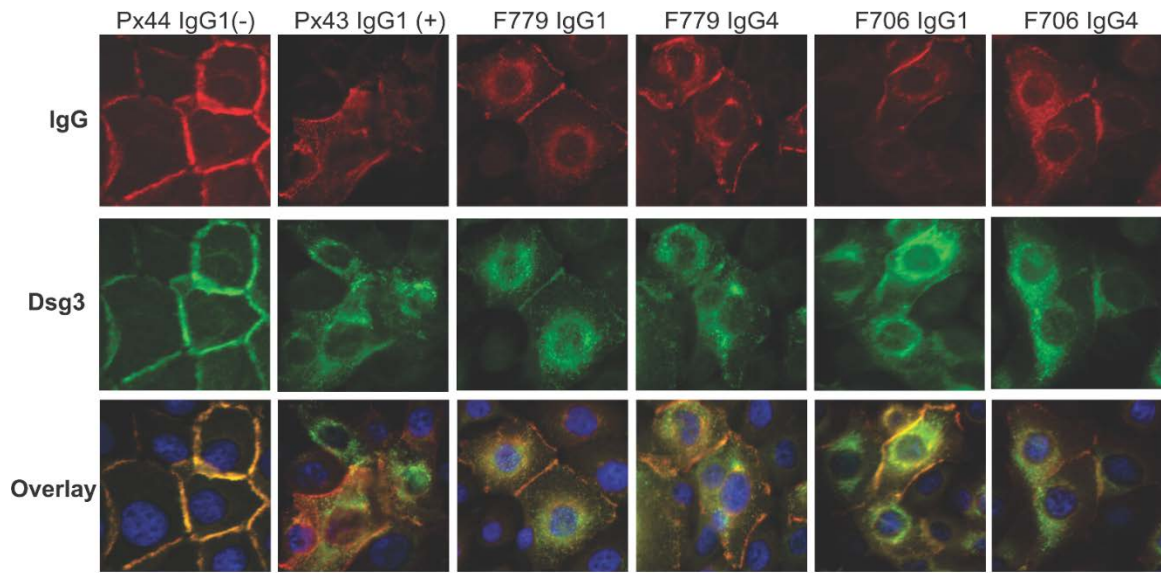


**Figure 2-2. Immunoreactivity of anti-Dsg3 PV mAbs by ELISA and IIF.** (A) Serial dilutions of recombinant human IgG1 ( $\blacklozenge$ ) or IgG4 ( $\square$ ) variants of mAbs (F706 against Dsg3 in upper left, F779 against Dsg3 in upper right, Px43 against Dsg3 in lower left and Px43 against Dsg1 in lower right) were incubated with Dsg3-coated ELISA plates. Bound antibody was detected using HRP conjugated mouse anti-human (H+L) Ab. Results are the mean of duplicate wells and are representative of 2-3 independent experiments. The relative affinity of each Ab was calculated by Lineweaver-Burk plot. (B) IgG1 and IgG4 variants of F706 and F779 mAbs show comparable immunoreactivity to the Dsg3+ human oropharynx tumor cell line SCC-47. F71 IgG1 and b12 IgG4 served as isotype controls. Other negative controls included HRP substrate OPD added alone or with secondary reagents only (secondary (2nd) biotinylated anti-human Fc Ab (Biotin Ab) or 2nd Ab with HRP-conjugated streptavidin Ab (Bio + SA Abs) in absence of tested PV mAbs. OPD: o-phenylenediamine dihydrochloride. (C) Px43 IgG1 and IgG4 demonstrate different relative affinities by ELISA but comparable affinity by indirect immunofluorescence (IIF) binding of human skin (red box). D: dermal part of skin. Results are representative of 1-2 experiments.





**Figure 2-3. IgG1 and IgG4 variants of anti-Dsg3 PV mAbs are pathogenic and cause suprabasal blisters in human skin.** (A) 25  $\mu$ g anti-Dsg3 PV IgG1 or IgG4 was injected ex vivo into human skin sections and incubated for 18 hours prior to harvest for histologic analysis. IgG1 and IgG4 of three anti-Dsg3 PV mAbs, F706, F779 and Px43 induced suprabasal blisters in human skin. PBS/ETA served as negative control. Scale bar, 100 microns. (B) The extent of histologic blistering caused by IgG1 and IgG4 mAb variants is comparable. (C) Direct immunofluorescence (DIF) staining of human skin with PV isotype mAbs illustrates cell surface binding in skin epidermis following injection of mAbs. D: dermal part of skin. Scale bar, 20 microns.



**Figure 2-4. Pathogenic F706 and F779 IgG1 and IgG4 cause loss of cell surface Dsg3.** Primary human keratinocytes were incubated for 8 hours with 50  $\mu\text{g}/\text{mL}$  F706/F779 IgG1 or IgG4. Cells were prepared for IF staining using anti-human IgG or anti-Dsg3 (5G11) primary antibodies. Px44 negative and Px43 positive control antibodies were previously characterized (386). Scale bar, 20 microns. Results are representative of two independent experiments.

## **CHAPTER 3:**

### **Lineage Analysis of Immune Repertoires in Pemphigus Vulgaris Reveals Distinct Pathways for Development of Autoreactive IgG4 versus IgG1 and IgA B Cells**

#### **3.1 Abstract**

Analyzing lineage relationships among autoreactive B lymphocytes is key to understanding the origins of autoimmunity. In pemphigus vulgaris (PV), characterized by blistering of the skin and mucosa, autoreactive anti-desmoglein (Dsg) antibodies predominantly belong to the IgG4 antibody subclass, and less prominently IgG1 and IgA. This compelled us to investigate whether anti-Dsg IgG4 B cells share lineages with anti-Dsg clones from the IgG1, IgA1, and IgA2 subclasses, suggestive of either evolution from common precursors or direct parent-child relationships. By sequencing B cell repertoires from the blood of four PV patients, we were able to sample over 70% of the circulating B cell lineages of each of subclass in each patient. We combined this approach with subclass-specific antibody phage display to identify and mark Dsg-reactive lineages within the larger repertoire. After identifying several Dsg-reactive lineages bridging subclasses, we found that the majority of anti-Dsg IgG4 B cells lack identifiable circulating relatives in other antibody subclasses. Specifically, direct IgG1-IgG4 switching comprises a minority pathway for the development of anti-Dsg IgG4 B cells, even when circulating IgG1 and IgG4 B cells appear to share a lineage. Within circulating anti-Dsg B cell lineages, parent-child relationships were observed between IgA1 and IgA2, between

IgG1 and IgA1, and between IgG4 and IgA2, indicating that IgA B cells in PV often derive either directly from or independent of IgG B cells and often switch between IgA1 and IgA2 subclasses. These findings suggest that class-switching between IgG and IgA may serve as a disease-modifying pathway in PV, and that the reciprocal relationship between IgG1 to IgG4 observed in chronic antigen stimulation may be in part due to the production of new antigen-specific lineages from non-IgG1 subclasses, rather than sequential class-switching of already extant IgG1 clones.

### **3.2 Introduction**

The staggering sequence diversity of antibody variable regions comes at a price – the generation of autoreactive clones alongside those required to protect from pathogens (140,395). While the variable region determines antigen reactivity, the antibody constant region confers diverse effector functions (e.g. agglutination, complement fixation, and activation of different leukocyte populations through Fc receptors) that can also contribute to the pathogenesis of autoimmunity by causing end-organ damage (396,397). During clonal expansion and affinity maturation, B cells can change their constant region to one of nine subclasses by recombining the IgH constant region locus – a tightly-regulated process known as class switching (388). While most current knowledge on human B cell class switch pathways derives from *in vitro* experiments (398,399), the advent of high-throughput sequencing has recently allowed the collection of large-scale *in vivo* data on class-switch relationships in normal and

allergic individuals (133)(137). However, very little is known about class switch pathways that contribute to human autoimmunity.

An ideal disease for investigating the potential role of class switch in autoimmunity, pemphigus vulgaris (PV) is a potentially fatal condition in which autoantibodies against the keratinocyte adhesion protein desmoglein 3 (Dsg3) induces mucosal blistering, with subsequent development of antibodies to Dsg1 causing wider skin involvement in its mucocutaneous subtype. Similar to other autoimmune diseases (242), PV patients with active disease demonstrate predominantly anti-Dsg IgG4 and less so anti-Dsg IgG1 (243); PV autoantibodies do not appear to associate with IgG2 and IgG3 subclasses (243,246). This anti-Dsg IgG4 predominance is so pronounced that PV patients demonstrate enrichment of total serum IgG4 (244). Intriguingly, whereas PV patients with active disease demonstrate anti-Dsg3 IgG1 and IgG4, unaffected relatives of PV patients (rarely), HLA-matched controls (rarely), and PV patients in remission can demonstrate anti-Dsg3 IgG1 (245,378). Th2-type cytokines such as interleukin-4 can stimulate B cell class switch to both IgG1 and IgG4 (400), although whether this switch occurs independently or sequentially is unknown. Additionally, the majority of PV patients demonstrate serum anti-Dsg IgA during active disease (243,258,259). Given that disease often begins in the mucosa and mucosal IgA plasmablasts can be self-reactive or polyreactive even in non-autoimmune subjects (265), IgA B cells could serve as an origin for autoreactive IgG4 clones.

The existence of multiple subclasses of autoantibodies targeting the same antigen suggests that class-switching is likely an important mechanism of pathogenesis

and diversification of the autoreactive B cell response in PV, and led us to investigate the origin of autoreactive B cells, particularly IgG4 B cells. To trace the developmental pathway of autoreactive B cells in PV across four key subclasses (IgA1, IgA2, IgG1, and IgG4), we paired deep sequencing of subclass-specific B cell repertoires with antibody phage display against Dsg to identify antigen-specific heavy chains, allowing us to simultaneously identify anti-Dsg clones and . This combined approach allowed us to both characterize the global lineage relationships between these four subclasses and to define the class-switch pathways that lead to the formation of autoreactive B cells in PV.

### **3.3 Materials and Methods**

#### **3.3.1 Peripheral Blood Collection and Processing**

Phage display libraries were produced as previously described (401), with minor modifications to generate biological replicates and facilitate the production of IgG1, IgG4, IgA1, and IgA2 specific amplicons for Next Generation Sequencing (NGS) *and* antibody phage display. In brief, 60 mLs of lithium-heparin anticoagulated peripheral blood was collected by venipuncture from four patients (PV8, PV16, PV17, PV18) following informed consent using an Institutional Review Board-approved protocol. Patient characteristics can be found in Table 3-1. The diagnosis of PV was established by typical clinical presentation, histology and immunofluorescence. For PV17 and PV18, two separate 30 mL blood samples were collected on the same day and processed independently to evaluate reproducibility among biological replicates.

Peripheral blood mononuclear cells (PBMCs,  $\sim 10^8$  per 60 mL blood) were isolated by Ficoll-Paque Plus (GE healthcare) gradient centrifugation according to the manufacturer's recommendations and snap-frozen in methanol prior to storage ( $-80^\circ\text{C}$ ) and further processing.

### **3.3.2 Library Preparation and Sequencing**

Total RNA was isolated from PBMCs after rotor-stator-mediated cell disruption/homogenization (Omni international) with the RNeasy Midi Kit (Qiagen) according to the manufacturer's recommendation. The typical yield from  $10^8$  PBMCs was 60-120ug of total RNA. Oligo-dT primed mRNA was reversely transcribed into cDNA in 15ug batches and subsequently destroyed by RNase digest with the Superscript First Strand synthesis system for RT-PCR (Thermo Fisher) following the manufacturer's protocol. In order to create NGS and phage display compatible amplicons, cDNA (equivalent of 7-8ug total RNA for each subclass) was subjected to 30 cycles of PCR with primers targeting the 5' end of the VH gene and the hinge region of the respective isotype subclass (subsequently called 1<sup>st</sup> round PCR reaction). The 1<sup>st</sup> round PCR reaction was conducted with Platinum Taq high fidelity polymerase (Thermo Fisher) as opposed to proof reading enzymes with lower error rate to maximize the yield and diversity of the libraries and all subsequent steps. PCR products were identified by gel electrophoresis, purified (Wizard SV PCR Clean-Up system, Promega) and reamplified in a semi-nested manner with primers for the 5' end of the VH gene and primers targeting the 5' end of the first constant region (CH1), leaving the first 9 nucleotides of this region

unprimed to allow for identification of the original template encoded isotype (i.e. IgA vs. IgG; this is subsequently called the 2<sup>nd</sup> round PCR reaction). The primer position at the 5' end of the CH1 harbors 2 distinct advantages that we reasoned were essential: (1) it allows the isotype to be directly confirmed from the template (as opposed to using J gene primers), while (2) it maximizes the length of the overlap during assembly of the paired end NGS reads, thereby creating higher confidence base calls in a 150 base pair window including CDR1, CDR2 and a significant proportion of the CDR3.

We generated phage display and NGS libraries from the same first round amplicon by modifying the flanking regions of the 2<sup>nd</sup> round PCR primers: In case of further processing for antibody phage display, the primers of the 2<sup>nd</sup> round PCR reaction were flanked by sequences encoding the GS-linker (5' primer; to be overlapped with the light chain amplicons for generation of scFvs) and a library-specific signature (3' primer, to exclude phage-mediated contamination during library screening).

NGS libraries were generated with 2<sup>nd</sup> round PCR primers that were flanked by a transposase sequence (5' and 3' primers) and in addition encoded a library-specific signature (3' primer only), thereby implementing a triple protection against falsely assigned cross-isotype relationships: the template-encoded isotype-specific start of the CH1 region, the library-specific signature, and the amplicon-specific Illumina barcodes. In order to minimize sampling error, 2<sup>nd</sup> round PCRs were performed with 3 and 6 technical replicates per primer combination for phage display libraries and NGS, respectively. Subsequent processing of the amplicons for phage display and NGS is described in the following 2 sections.



### 3.3.3 Antibody Phage Display

To identify Dsg-reactive heavy chain sequences in the deep sequencing data, antibody phage display libraries were constructed using the aforementioned isotype-specific VH amplicons generated during 30 cycles of 2<sup>nd</sup> round PCRs. For PV17 and PV18, 1<sup>st</sup> round PCR products were produced for each biological replicate and subjected to 2<sup>nd</sup> round PCR amplification after equimolar pooling. These VH amplicons were fused by overlap extension PCR with VL amplicons to create scFv sequences with the configuration VL-GS-linker-VH. The scFvs sequences were flanked by SfiI sites (introduced during overlap extension PCR), which allows for directional cloning into pComb3X as previously described (208). Electroporation (Micropulser Electroporation apparatus, Bio-Rad,  $12.5 \text{ kVcm}^{-1}$ , exponential decay with time constants between 3.4-4.0ms) of libraries into XL-1 blue cells (Agilent) achieved a total number of  $10^8$ - $10^9$  transformants per library. Libraries were generated and screened separately. All libraries were panned separately against Dsg3 and Dsg1 at room temperature for three to five rounds of panning, with the exception of PV16 IgG4 and PV16 IgG1, which were panned against Dsg3 and against Dsg3 and Dsg1 in alternation (in the order 3,1,3,1). Up to 200 colonies were analyzed for each patient from all rounds of panning that showed binding to Dsg1 or Dsg3 by pooled phage ELISA.

Unique clones were produced as soluble scFv in *E. coli*. In brief, 2ul of plasmid was added to 5-10 ul of Top10F' cells, then transformed and plated on LB/ampicillin plates according to the manufacturer's instructions. Single colonies were diluted into 20mL SB containing 20mM MgCl<sub>2</sub> and 0.05 mg/mL carbenicillin, grown overnight at 30C,

then diluted 1:100 in 400mL of SB with the same concentration of MgCl<sub>2</sub> and carbenicillin. Cultures were grown to OD<sub>600</sub> of 0.5-1.0, then induced with 1mM IPTG and grown at 30C for an additional four hours (which was further optimized for some clones to improve production). Cultures were spun down and pellets were re-suspended in 30-35 mL of FastBreak lysis reagent (Promega), then processed and subject to cobalt affinity chromatography (Talon, Clontech) according to the manufacturer's instructions. Purified scFv concentration was determined by SDS-PAGE followed by Coomassie blue staining, using BSA standards for quantitation.

Cloned scFv were validated using indirect immunofluorescence at 50 mg/mL on frozen normal human skin sections (obtained through the Penn Skin Disease Research Center) in TBS buffer containing 1mM CaCl<sub>2</sub> and 1% BSA. Bound antibody was detected with rat anti-HA antibody (3F10, Roche), then by 1:200 Alexa Flour 594 donkey anti-rat IgG (Thermo Fischer Scientific). IF was visualized using an Olympus BX61 microscope, images acquired using SlideBook 4.2 software (Olympus) and a Hamamatsu Orca ER camera.

Cloned scFv were also tested by ELISA against Dsg3, Dsg1, and an irrelevant antigen (either tetanus toxoid or BP180) as a negative control (Euroimmun), performed according to manufacturer's instructions.

### **3.3.4 Next Generation Sequencing**

Isotype-specific 1<sup>st</sup> round PCR VH amplicons were amplified with 2<sup>nd</sup> round PCR primers as mentioned above. In order to minimize PCR induced mutations, the NGS

libraries were produced with Q5 polymerase (NEB) using a maximum of 3ng per reaction and only 12 PCR cycles. After visualization and purification (Wizard SV PCR Clean-Up system, Promega), the 2<sup>nd</sup> round product concentration was quantified (NanoDrop, Thermo Fisher) and PCR products were pooled with equal amounts of DNA input representing each primer pair. Subsequently, pooled PCR products were subjected to 12 cycles of PCR amplification to attach the Illumina Nextera barcodes and adapters, followed by gel purification (Wizard SV gel purification system, Promega) and additional purification with AmPure beads (Beckman Coulter) according to the manufacturer's recommendations. NGS library concentrations were determined (Qubit dsDNA HS, Thermo Fisher) and pooled accordingly. The concentration of sequencing-competent amplicons was additionally analyzed by qPCR (KAPA Biosystems), and the purity was confirmed (Bioanalyzer, Agilent) before subjecting the amplicons to paired end 2x300 base pair sequencing (Illumina MiSeq). In order to estimate the frequency of artificial mutations introduced during library production and sequencing, we generated an additional amplicon with comparable length derived from a Sanger sequencing-verified plasmid encoding the CD8-CD137 transition of a chimeric antigen receptor (CAR). The plasmid was subjected to the same number of PCR cycles and procedures as the abovementioned amplicons and added to the library mixture comprising 1% of the total amount of DNA (as determined by Qubit).

Technical replicates were generated by production of 2 NGS libraries from the same cDNA with identical primers (except for barcodes and signatures) and sequencing those 2 libraries separately. Biological replicates were generated by processing 2 blood

samples (so that B cell clones shared between the replicates have to be derived from each aliquot and therefore represent physically separate B cells) separately for NGS library production.

### **3.3.5 Raw Data Processing and Lineage Clustering**

Libraries were sequenced as mentioned above using Illumina Miseq 300bp paired end sequencing (Institute for Genome Science, University of Maryland; Next-Generation Sequencing Core, University of Pennsylvania). Forward and reverse reads were merged using PEAR (402), and identical reads were merged into non-redundant (NR) sequences. All non-redundant sequences with isosignature exactly matching the predicted isosignature for a given library were kept, and uploaded to IMGT-Hi-VQUEST for VDJ assignment. The IMGT results were filtered to only productive hits.

Clonality was defined by first binning all sequences from all four patients by V, J and CDR3 length, then for each bin calculating the nearest neighbor nucleotide hamming distance from a given CDR3 to all CDR3 in the same V-J-length bin belonging to another patient. This allowed us to define an upper limit on allowed clonal divergence. A per-patient nucleotide hamming distance cutoff was selected such that the resulting clusters would have 99% confidence of unrelated sequences being excluded. Sequences within each patient-subclass pair were then clustered using a customized perl script and a centroid (UPGMC) linkage function implemented in the Algorithm::Cluster module. Within each patient-subclass pair, clusters were subsequently superclustered by their consensus CDR3 sequence, without regard to VDJ gene assignment, and only the cluster

with the largest number of reads (toprank) was considered for large scale subclass connectivity analysis. Heavy chains found by phage display were matched back to NGS clusters by searching for consensus CDR3 falling within the original clonality CDR3 nucleotide identity cutoff for each patient.

### 3.3.6 Estimation of Clonal Richness/Diversity and Total Population Size

To estimate the total population size and the fraction of the total population that was captured by sequencing, all biological and technical replicates for a given subclass-patient pair were pooled and subject to rarefaction analysis and asymptotic diversity estimation using the iNEXT package in R (304) and the online implementation of iNEXT (<http://chao.shinyapps.io/iNEXTOnline/>). Each supercluster was treated as a “species”, and abundance-based rarefaction was conducted on the number of reads. Furthermore, in order to determine the amount of total diversity in the sample capture, the asymptotic Shannon Diversity, Simpson Index, and Species Richness were compared to the actual values for those metrics in a given sample. To correct for the artificial inflation of low-abundance taxa due to sequencing and amplification error, the number of singletons in each sample was deflated using the formula presented in (309):

$$\hat{f}_1 = \frac{2f_2^2}{3f_3} + 2f_2 \left( \frac{f_2}{3f_3} - \frac{f_3}{4f_4} \right)$$

To determine the reproducibility of lineages among technical and biological replicates, capture-recapture analysis was performed using the Chapman estimator (298):

$$N_C = \frac{(n_1+1)(n_2+1)}{(n_{12}+1)}$$

$N_C$  = population size

$n_1$  = number of lineages in first replicate

$n_2$  = number of lineages in second replicate

$n_{12}$  = number of lineages in both replicates

A lineage was considered existing in one replicate if it appears with five reads in that replicate, and existing in both replicates if it appears with at least five reads in each replicate in question.

### 3.3.7 Quantitation of Subclass Connectivity

Average pairing rates between subclasses were computed as the harmonic mean of the frequency of shared lineages appearing between the subclasses, in order to normalize for differing subclass sizes. For example, if  $n$  lineages are shared between subclasses of read size  $A$  and  $B$ , the pairing rate is

$$\frac{2}{\frac{A}{n} + \frac{B}{n}}$$

For comparing pairing rates between the anti-Dsg and global repertoire, the calculated pairing rate was bootstrapped down to the size of the anti-Dsg repertoire for any given subclass pair.

### **3.3.8 Lineage Analysis**

Lineages of interest were selected as those phage-display-derived heavy chain sequences mapping to top-rank clusters in other subclasses with at least 5 reads, with an additional constraint that the top-rank cluster in the “home” subclass of the Dsg-reactive clone have an identical V-gene (without consideration of the allele). Sequences within each lineage of interest were subject to primer trimming and full-length clustering at 97% radius using the `cluster_otus` command in USEARCH (315). Trees were drawn after removing all nodes with fewer than 5 reads. For visualization, some trees were drawn with all nodes with fewer than 10 reads removed, though this did not affect their topology. Parameters were tested on a monoclonal CD137 template subject to the same amplification procedure as the libraries (Figure 3-23). In total, 19 trees were produced (Table 3-2). Maximum parsimony trees were generated using the Change-O package (116), which calls `dnapars.exe` from PHYLIP (403). Intermediates from maximum parsimony trees were synthesized as gblock gene fragments (Integrated DNA Technologies) with nucleotide additions allowing either overlap PCR with the corresponding light chain or direct cloning into the pComb3X vector.

### **3.3.9 Epitope Mapping**

Epitope mapping was performed as previously described for patient sera (22) and scFv (24), using chimeric Dsg3/2 or Dsg1/2 molecules expressed in either baculoviral or HEK293-derived supernatants. Dsg3-reactive clones were tested by immunoprecipitation against Dsg3/2 constructs followed by western blotting using an anti-E-tag HRP antibody, and Dsg1-reactive clones were tested against Dsg1/2 constructs in an analogous manner

## **3.4 Results**

### **3.4.1 High-Throughput Sequencing of Subclass-Specific B Cell Repertoires in PV Indicates a Predominance Of Anti-Dsg IgG4 Lineages, Followed by IgA1, IgG1, and IgA2**

A better understanding of the isotype relationships of autoreactive B cells requires deep sequencing of B cell repertoires to obtain large-scale B cell lineage information, as well as a method for identification of antigen-specific clones. We performed high-throughput sequencing of IgG4, IgG1, IgA1, and IgA2 B cell repertoires in a panel of four PV patients with active disease (patient characteristics can be found in Table 3-1). RNA was isolated from PBMC derived from 50-60 mL of peripheral blood, representing approximately  $2 \times 10^6$  class-switched B cells ( $[50 \text{ mL of blood}] \times [4 \times 10^5 \text{ B cells/mL of blood}] \times [10\% \text{ B cells are class-switched}]$ ), per patient. Sequencing libraries



were prepared from an amplicon generated using a 3' primer specific for the unique hinge region of each subclass. Subsequent amplification for library production was conducted using a 3' primer containing a constant-region-specific binding sequence, a custom library barcode for every patient-subclass pair, and an Illumina index sequence. The constant region sequence and the library barcode were collectively called the isosignature (Figure 3-1). Ensuring the index sequence and isosignature were concordant allowed high-confidence discrimination between subclasses and stringent filtering of potentially mis-assigned sequences that would be falsely interpreted as cross-subclass relationships. This amplification scheme allowed us to generate ~225,000 to ~6 million paired-end raw reads per patient-subclass pair on the Illumina MiSeq 2x300 bp platform (Table 3-3). Furthermore, each sequencing run was accompanied by a monoclonal template subject to the same amplification steps, to give a bound on amplification and sequencing-induced error.

Antibody phage display libraries were prepared from the same hinge-primed amplicon and panned against Dsg3 and/or Dsg1 to select subclass-specific Dsg-reactive clones. Across 16 libraries representing four patients and four subclasses, each representing 100 million to 1 billion total transformants, up to 200 clones were characterized for anti-Dsg ELISA reactivity and cell-surface staining by indirect immunofluorescence to independently verify the PV phenotype, resulting in 96 total unique anti-Dsg B cell clones representing 80 lineages (Table 3-4). Characterized clones showed a wide variety of V-gene usage and a lack of stereotyped CDR3 sequences, demonstrating a highly polyclonal anti-Dsg response across subclasses. The plurality of

unique Dsg-reactive clones belong to the IgG4 subclass (34), followed by 21 IgA1 clones, 18 IgG1 clones, and 13 IgA2 clones. We found a comparable number of IgA and IgG4 clones among the four patients (34 of each), suggesting a diversified anti-Dsg response in both the IgG and IgA isotypes.

### **3.4.2 Anti-Dsg Clonal Lineages Bridge Multiple Subclasses**

After sampling anti-Dsg clones from each of the sixteen subclass-patient libraries, we first sought to determine the lineage relationships of these Dsg-reactive clones across the four sequenced subclasses. To that end, we determined the clonal structure of the global antibody repertoire by partitioning sequences into lineages. First, across all patients, sequences sharing V, J, and CDR3 length were binned together. Then, by comparing the CDR3 sequences from each of the four patients against the other three, we derived a per-patient CDR3 nucleotide identity cutoff that gives 99% confidence that two clustered sequences do not belong to different lineages (Figure 3-2). CDR3 nucleotide similarity cutoffs ranged from 80.6% in PV8 to 86.1% in PV18. Finally, to mitigate the effects of VJ misassignment and sequencing error creating spurious lineages, we further grouped all clusters together that meet the same CDR3 cutoff, without requiring matching V or J genes; this allowed us to sort clusters by read count into ranked lists called superclusters, and to eliminate non-top-ranked clusters within a supercluster from consideration during lineage analysis to improve stringency.

We subsequently mapped the phage-display derived heavy chain sequences to deep-sequencing derived clonal lineages, finding 80 Dsg-reactive lineages that span between one and four subclasses (Figure 3-3a). We further analyzed these 80 Dsg-reactive lineages in terms of pairwise overlaps between subclasses (Figure 3-3b and 3-3c), which demonstrated that a majority of IgA1- and IgA2-containing Dsg-reactive lineages contain relatives in other subclasses; in particular, the majority of IgA1 have relatives in IgA2 and vice versa. By contrast, only a minority of IgG4-containing Dsg-reactive lineages can be found in other subclasses. It is noteworthy that when IgG4-containing lineages bridge subclasses, that they show a slight preference for IgG1 (21%) over IgA1 (15%) and IgA2 (13%). To determine whether IgG4 is unique among subclasses in its lack of relationships, we compared the proportion of single-subclass and multi-subclass Dsg-reactive lineages between all four subclasses (Figure 3-3d). Only 28% of IgG4-containing Dsg-reactive lineages are multi-subclass, which, by Fisher's exact test, is significantly fewer than the other three subclasses combined ( $P = 6.31 \times 10^{-6}$ ).

### **3.4.3 Large-Scale Clonal Analysis Indicates anti-Dsg IgG4 B cells Show Increased Connections to Other Subclasses Relative to IgG4 B cells in the Global Repertoire**

We have demonstrated that the anti-Dsg IgG4 repertoire is disconnected from other subclasses compared to anti-Dsg lineages in other subclasses, and that the anti-Dsg IgA repertoire is highly connected. To determine whether this is a unique feature of

the anti-Dsg repertoire relative to the global B cell repertoire, we next sought to compare the connectivity of the anti-Dsg B cell repertoire to the global repertoire. To understand the large-scale structure of the repertoire, we calculated the average pairing rate of each subclass pair by taking the harmonic mean of the proportion of overlapping lineages between a given subclass pair (see Methods). First, we compared the between-subclass pairing rate of the Dsg-reactive and global repertoire for all four patients and all four subclasses. This demonstrated that the Dsg-reactive repertoire is significantly more interconnected than the global one ( $p < 0.0001$ , Figure 3-4a). Separating out the Dsg-reactive repertoire by subclass shows that IgA1 and IgA2 show the highest interconnectedness, both in the global and Dsg-reactive repertoires (Figure 3-4b).

We next determined the subclass pairings which were enriched in the Dsg-reactive repertoire in order to determine lineage relationships important in PV. For each of the six possible subclass pairs, we normalized the pairing rate of the Dsg-reactive repertoire to the global repertoire (Figure 3-4c). To account for the possibility that this increase in IgG4 connectivity is simply a positive correlation with lineage size, we compared the number of reads mapping to the IgG4 subclass in single-subclass anti-Dsg lineages vs multi-subclass anti-Dsg lineages containing IgG4 in each patient. No significant difference in IgG4 read number between multi-subclass and single-subclass autoreactive lineages was found, indicating that lineage size and/or sampling depth appears to have a minimal effect on connectivity (Figure 3-4d). Though it remains isolated in an absolute sense, the Dsg-reactive IgG4 repertoire shows an increased connectivity to other subclasses compared to anti-Dsg lineages in the global repertoire.

### **3.4.4 Anti-Dsg IgG4 B Cells Predominantly Demonstrate No Subclass Connectivity but can Rarely Switch Directly From anti-Dsg IgG1 B Cells**

Our global repertoire data demonstrates that anti-Dsg IgG4 is disconnected from the larger anti-Dsg repertoire relative to other subclasses. Of the 39 total IgG4-containing Dsg-reactive lineages, 28 contained sequences from only IgG4, while 11 contained sequences from multiple subclasses (Figure 3-3d). To better understand the clonal development of the Dsg-reactive IgG4 repertoire on a fine level, we further analyzed seven lineages containing both IgG4 and IgG1 by generating maximum parsimony phylogenetic trees for each lineage (Figure 3-5, 3-6, 3-7, 3-8). 1/7 of these lineages showed evidence of a direct IgG1 to IgG4 switch, with the IgG1 clone differing from the IgG4 clone by a single substitution (Figure 3-5a). The remainder do not show an identifiable direct switch between IgG1 and IgG4, either due to very early divergence from germline and subsequent independent evolution of each subclass (Figure 3-5b, 3-7), late divergence after some mutation of a precursor (Figure 3-5c), or show co-evolution and intermingling between IgG1 and IgG4 without evidence of direct class switch (Figure 3-5d).

To understand the onset and development of Dsg reactivity within each lineage, we produced several selected sequenced and inferred nodes from seven of the IgG4-containing multi-subclass lineages and several sequences with a germline-reverted heavy chain from lineages containing only IgG4 Dsg3. ELISA indicated that in all except one multi-subclass lineages tested, relatives of Dsg-reactive clones in other subclasses

showed similar ELISA reactivity (Figure 3-21a); only the PV17.G1.D1.P4.8 lineage (Figure 3-5d) showed no reactivity in other subclasses. Furthermore, in multi-subclass lineages with confirmed reactive sequences across multiple subclasses, 4/6 most recent common ancestors (MRCA) and 8/11 germline heavy chain sequences (Figure 3-21b) show strong anti-Dsg reactivity mirroring the reactivity of phage-display derived clones and sequenced nodes.

### **3.4.5 Anti-Dsg IgG4 B Cells Evolve Independently of IgA1 but can Directly Give Rise to anti-Dsg IgA2 B Cells**

As PV blistering often begins in the mucosa, we next sought to identify potential direct-switch relationships between IgA clones and IgG4 that could imply a mucosal origin for autoimmune B cells in PV. Anti-Dsg3 IgG4 clones share three lineages with IgA2 and two with IgA1. IgA1 and IgG4 show no evidence of direct switching, lying on different branches of the two analyzed lineages in which they coexist (Figure 3-5a, 3-5d). Two IgG4-IgA2 lineages show a direct switch from IgG4 from IgA2 (Figure 3-8, 3-9), while in the other case, the IgA2 switches from IgA1 and the IgG4 switches from IgG1 (Figure 3-5a). In both of these IgG4-IgA2 direct switches, the IgA2 clone remains anti-Dsg reactive (Figure 3-21a).

### **3.4.6 Anti-Dsg IgA1 Shows Close Relationships to Both anti-Dsg IgG1 B Cells and anti-Dsg IgA2 by Direct Switch**

After studying the relationships between IgG4 and each of the other three subclasses, we next sought to apply the same analysis to the lineage relationships between IgA1, IgA2, and IgG1. Interestingly, IgG1 and IgA1 show a direct switch in four of six analyzed lineages in which they coexist. We also found that in two lineages demonstrating an IgG1 to IgA1 switch, there is evidence of a subsequent switch to IgA2 with a few mutations (Figure 3-9, 3-10), suggesting that IgA1 can act as a hub for clonal maturation from IgG1 to IgA2. There are no examples of direct IgG1 to IgA2 switching present without IgA1 sharing the IgG1 sequence, implying the necessity of an IgA1 intermediate in this switch pathway.

The connection between IgA1 and IgA2 is particularly strong in the autoreactive lineages we have studied, mirroring their high pairing rate in the global repertoire. There is evidence of direct switch in 10/11 analyzed lineages in which IgA1 and IgA2 coexist. In the other case (Figure 3-16), the IgA1 and IgA2 appear on adjacent terminal nodes, but show no direct switch. In total, of the 13 multi-subclass lineages containing IgA2, 10 show a direct switch from IgA1 (two of which originate from a shared IgG1/IgA1 node), two show a direct switch from IgG4, and another demonstrates co-evolution with IgA1.

### 3.4.7 Repertoire Sequencing Effectively Captures Each Subclass' Diversity

Our study indicates that the majority of anti-Dsg IgG4 B cells in PV do not demonstrate lineage relationships to IgG1 or IgA. Our inability to identify IgG1-IgG4 B cell lineage relationships or direct class-switches could be due to the fact that such relationships simply do not occur, meaning IgG4 B cells are switching directly from upstream non-IgG1 subclasses (IgM, IgG2, IgG3). Alternatively, we simply could have failed to sample the IgG1 precursor of any given IgG4. Because we are able to identify cross-subclass lineage relationships and ongoing class-switching for the majority of anti-Dsg IgA1 B cell lineages, a population that is likely as diverse in the circulation as IgG1, and because we are able to find several instances of ongoing IgG1 to IgA1 class-switching, we find it unlikely that we lack the power to detect relevant class-switching relationships between IgG1 and IgG4 due to undersampling. However, to address the issue of potential undersampling of the circulating repertoire, we performed rarefaction analysis and asymptotic estimation of sample coverage for each patient-subclass pair (Figure 3-22). Asymptotic estimation of species richness (e.g. determining the total number of lineages in the population by estimating how many lineages we would find had we continued sampling indefinitely) shows that we have captured 71-97% total estimated lineages present in each patient. Furthermore, we also determine the percentage of total lineage diversity we have sampled using asymptotic estimation of the Shannon diversity and Simpson index, which are used in ecology to determine the diversity of common and dominant taxa in ecological samples. Looking across each patient-subclass pair, we have captured 97-99.98% of the lineage Shannon diversity and



99.6-100% of the lineage Simpson diversity, indicating we have sampled most of the common and dominant antibody sequences from circulating IgA1, IgA2, IgG1, and IgG4 B cells in the repertoire.

To determine the relative contribution of technical and biological factors to undersampling, biological replicates were performed for PV17 and PV18 by taking two separate 30 mL blood draws, and for PV16, PV17, and PV18, each biological replicate (one for PV16, two for PV17 and PV18) was sequenced twice as technical replicates to determine the sufficiency of sequencing depth. Each pair of biological or technical replicates were compared for overlapping lineages, and population size estimators were used to determine what fraction of lineages were being captured (Table 3-5). After filtering those lineages containing below five reads, mark-and-recapture analysis using the Chapman estimator demonstrates that 96.57-99.97% of the lineages within a sample are captured by pooled technical replicates, and over 35.46-81.99% of the lineages within the circulation are captured by pooled biological replicates, suggesting that most of the undersampling in the experiment is due to biological, rather than technical, limits.

### **3.4.8 Epitope Mapping Shows N-terminal Preference of IgG4 Relative to Other Subclasses**

Our data have shown, with high confidence, that the anti-Dsg IgG4 repertoire does not share lineage relationships or direct class-switch origins from IgG1 or IgA1. We

have also demonstrated that the anti-Dsg IgG1, IgA1, and IgA2 repertoires often share lineages separate from anti-Dsg IgG4. To determine whether this division in the repertoire could be explained by differing epitope reactivity, we determined the Dsg extracellular (EC) domain specificity of Dsg-reactive clones (Table 3-7)(226). Dsg-reactive clones sharing a lineage mapped to the same epitope in all cases. The plurality of clones within each subclasses demonstrated reactivity against the amino-terminal EC1, and in aggregate, at least two Dsg-reactive clones react to each of the five extracellular domains. However, there is a statistically significant shift towards EC1 reactivity in the IgG4 subclass compared to all other subclasses ( $p = 0.009$  by Fischer's exact test). This is especially pronounced when comparing IgG4 to the IgA subclasses, which show more evenly-distributed reactivity to all five extracellular domains.

### **3.5 Discussion**

Class-switching is largely directed by the cytokine milieu and T-cell help that a B cell encounters during an ongoing immune response, where it serves to modify and diversify the effector functions of antibodies. In this study, we have presented evidence that despite co-occurring with several other subclasses and responding to the same antigen, the IgG4 Dsg-reactive repertoire arises largely independently of the other subclasses in a panel of PV patients. By cloning each subclass into a separate phage display library, we have the ability to screen for rare autoreactive clones from more infrequent B cell subclasses (like IgG4 and IgA2) at unprecedented depth, free of competition from clones in more diverse subclasses. Furthermore, our sequencing

method managed to capture a majority of the total lineages in each sample, including in particularly diverse subclasses like IgA1 and IgG1, allowing us to obtain a representative picture of the class-switch landscape in the circulating B cell repertoire of PV patients.

Only a minority of detected Dsg-reactive IgG4 lineages also contain IgG1 sequences, with only a single lineage showing a direct switch between IgG1 and IgG4. The existence of a switch pathway from IgG1 to IgG4 in the normal B cell repertoire was characterized in a large-scale repertoire study of the human class-switching landscape, with direct switches from IgM, IgG2, and IgG3 to IgG4 occurring at similar frequencies (137). Due to their relative unimportance in PV, our study did not include IgG2 or IgG3, leaving open the possibility of an IgG2 or IgG3 clone – most likely a non-reactive one – during a switch to IgG4. Even given that caveat, our data suggest a mostly disjoint relationship between the IgG1 and IgG4 anti-Dsg repertoires, in which both evolve independently from different precursors and with IgG4 only rarely switching from IgG1. Furthermore, only 28% of IgG4 belonged to multi-subclass lineages, and 4/6 lineages containing confirmed reactive IgG1 and IgG4 clones showed early divergence between the two, suggesting that even when anti-Dsg IgG1 and IgG4 share a lineage, they develop largely independently within said lineage.

Sequences mapped to anti-Dsg IgG4 lineages showed significantly fewer connections to other subclasses than IgA1, IgA2, or IgG1, suggesting a pathway in which the majority of the Dsg-reactive IgG4 repertoire originates from a non-IgG1 subclass (like IgM, IgG2, or IgG3), creating its own self-contained thoroughfare of autoreactivity in the overall class-switching landscape. This is further supported by the fact that a

majority of anti-Dsg3 IgG4 reverted to either their MRCA (2/3) or germline (8/11) sequence remained reactive. A similar finding has recently been published for an endemic form of pemphigus foliaceus found in Brazil (fogo selvagem); reversion of mutations to germline in two clones found both of them to remain cross-reactive to Dsg1 and the sandfly salivary antigen LJM11 (404). We have also previously found some anti-Dsg3 IgG using the VH1-46 gene segment retain reactivity at germline (367). Of the 14 anti-Dsg3 IgG4 clones we have reverted, only one used VH1-46 (though it was germline-reactive), demonstrating that the IgM repertoire may demonstrate a wider range of Dsg reactivity than previously suggested. Further studies, including repertoire cloning of IgM, IgG2, and IgG3, comparing germline-derived sequences from different subclasses in the anti-Dsg repertoire in terms of polyreactivity, affinity, and epitope preferences, and determining the integrity of central tolerance mechanisms in PV patients, will shed more light on the divergent origins of Dsg-reactive IgG1 and IgG4.

Interestingly, though anti-Dsg IgG4 was the most solitary of all the anti-Dsg subclasses, it demonstrated a noticeable increase in connectivity relative to the global anti-Dsg repertoire – in fact, this increase in connectivity between subclasses in Dsg-reactive lineages relative to the global repertoire occurs in all subclasses. Determining whether this increase in connectivity occurs in any ongoing antigen-specific immune response, or specifically in the context of PV or other autoantibody-mediated disease, will require further study. Theoretically, an increase in subclass connectivity in an antigen-reactive section of the B cell repertoire could simply reflect increased class-switching as part of an ongoing immune response. In the specific case of IgG1 and IgG4,

however, this connectivity is not the result of direct switching between the two subclasses, but due to lineages containing Dsg-reactive IgG1 and IgG4 sharing divergent origins from precursors. Conversely, the connectivity in anti-Dsg IgA1-IgA2 shared lineages seem to arise mainly from direct switching. This suggests that increased direct switching is not the only mechanism increasing subclass connectivity in an ongoing immune response. For example, it may be the case that IgM B cells with an intrinsic propensity for switching divergently to different subclasses are specifically induced to proliferate during PV, or that the cytokine milieu in PV specifically induces some IgM B cells to switch to IgG1 and IgG4 independently, thus driving an increase in subclass connectivity without an increase in direct-switch relationships.

The lack of identified precursor-product relationships between IgG1 and IgG4 may also be explained by the difficulty of capturing the full diversity of the human B cell repertoire. Through mark-and-recapture analysis, we have shown that our approach is primarily constrained by biological sampling. Although we used over 50 mL of blood (approximately 1% of total blood) from each patient to maximize coverage, we are still most likely missing rare circulating clones. This is made worse by the fact we may be biasing against sequencing memory B cell clones due to the presence of highly transcriptionally active plasmablasts “washing out” the mRNA contribution from memory cells. Furthermore, there is the possibility of non-circulating related IgG1 B cell clones occupying niches inaccessible to our sampling, or that hypothetical precursor IgG1 are completely consumed by the switch to IgG4 and have disappeared from the circulation by the time of disease onset. These caveats can be partly minimized due to

our rarefaction analysis showing that we have effectively captured over 70% of lineages in every sample, even in diverse subclasses, and by the fact that that we have found robust evidence of direct switching between IgG1-IgA1, IgA1-IgA2, and IgG4-IgA2 as part of an ongoing anti-Dsg response, making the absence of IgG1-IgG4 direct sequences particularly conspicuous in this context, even given the constraints listed. Nevertheless, given these caveats, it is entirely possible that longitudinal subclass-specific repertoire studies in PV, or studies of the splenic, mucosal, and/or bone marrow repertoire in PV, may reveal new relationships between subclasses that we are unable to detect in peripheral blood.

The class-switch pathways leading to IgG4 are of fundamental interest not only in the study of PV, but also the mechanism of several other autoimmune diseases (242) and states of chronic antigen stimulation. Beekeepers (248), patients with chronic helminth infections (252), and patients undergoing allergen desensitization therapy (251) all develop IgG4, which serves as a blocking antibody to dampen the response to the given antigen. In particular, novice beekeepers tend to show IgG1 to bee venom antigens, while more experienced beekeepers who have been tolerized to bee stings over time show IgG4 to the same antigens (256). Similarly, in fogo selvagem, healthy people in the endemic focus often display anti-Dsg1 IgG1, whereas patients with active disease display anti-Dsg3 IgG4 (405,406). A larger implication of the existence of independent IgG1 and IgG4 anti-Dsg repertoires is that the reciprocal relationship between IgG1 and IgG4 in some forms of chronic antigen stimulation may be due to separate populations of B cells independently switching to IgG1 and IgG4 in response to

antigen challenge, rather than direct switch between IgG1 and IgG4. It has been shown that in allergen desensitization therapy and in tolerized beekeepers, the blocking IgG4 is exclusively secreted by a population of IL-10 secreting regulatory B cells (407).

Expansions of regulatory B cells can be found in PV and other autoimmune conditions (408,409); while we cannot draw conclusions about cell of origin from our experiments, these data are which is mildly suggestive of the idea that pathogenic IgG4 in PV – and protective IgG4 in other forms of chronic antigen stimulation - may arise from these same cells developing independently of IgG1 responses to the same antigen. Single B cell cloning experiments in PV patients targeting this population, or similar experiments conducted longitudinally over a period of chronic antigen stimulation, will be able to shed light on this hypothesis. Determining the developmental pathway of pathogenic or protective IgG4 would lead to a better understanding of IgG4-mediated autoimmune disease, allergen tolerization, and other IgG4-dependent processes.

We have also characterized, for the first time, the lineage relationships of IgA in PV. We discovered a highly diverse IgA response to Dsg1 and Dsg3, which in several lineages overlaps strongly with the IgG1 repertoire by direct switching from IgG1 to IgA1. A high-throughput analysis of class switching recently showed that IgG1-IgA1 parent-child relationships are the second most common in the class-switching landscape (137). Both the direct IgM to IgA1 switch and the sequential IgM to IgG1 to IgA1 have previously been shown to be dependent on TGF-beta and CD40 activation in cultured human B cells (410). IgA switching can also be triggered in a T-cell independent manner on mucosal surfaces by TLR4 engagement and secretion of the cytokines BAFF and APRIL

by dendritic cells (264). The role of TGF-beta in PV is not well-characterized; one study found its serum levels elevated relative to controls (36), though others found otherwise (411,412). Serum BAFF and APRIL also have an unclear role, having been found to be elevated in bullous pemphigoid, but not PV (413,414).

Furthermore, we have characterized antigen-specific switching from IgG4 to IgA2 for the first time. The functional relevance of this switch pathway is unknown. 78% of all IgA2-containing Dsg-reactive lineages were multi-subclass, and all except one of the 12 well-represented IgA2 multi-subclass lineage characterized shows a direct switch from either IgA1 or IgG4, and the other lies close enough to IgA1 terminal nodes within the tree to suggest that the an IgA1 to IgA2 switch probably occurred but wasn't captured by sequencing. IgA2 is the predominant IgA subclass in the mucosa (415). The production of IgA2 is known to occur in the colonic lamina propria through secretion of IL-10 and APRIL secreted from colonic epithelium in response to normal intestinal flora; this occurs in both an IgM-to-IgA2 and IgA1-to-IgA2 fashion (91).

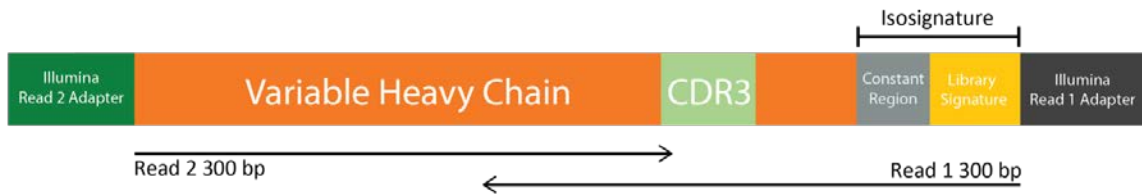
Given our results suggesting that anti-Dsg IgA clones either develop independently from, or switch directly from, anti-Dsg IgG clones, and literature showing that IgA-inducing cytokines have an unclear role in PV pathogenesis, it may be the case that the existence of anti-Dsg IgA1 and IgA2 is simply a post-hoc epiphenomenon of PV, catalyzed by the localized effect of IgA class-switch mediators like TGF-beta, LPS, BAFF, or APRIL on Dsg-reactive IgM, IgG1, or IgG4. This seems especially likely for IgA2, as it lies on the 3' end of the IgH locus and therefore cannot serve as a class-switch intermediate or harbor of pathogenic clones. It may be the case that the clones found in



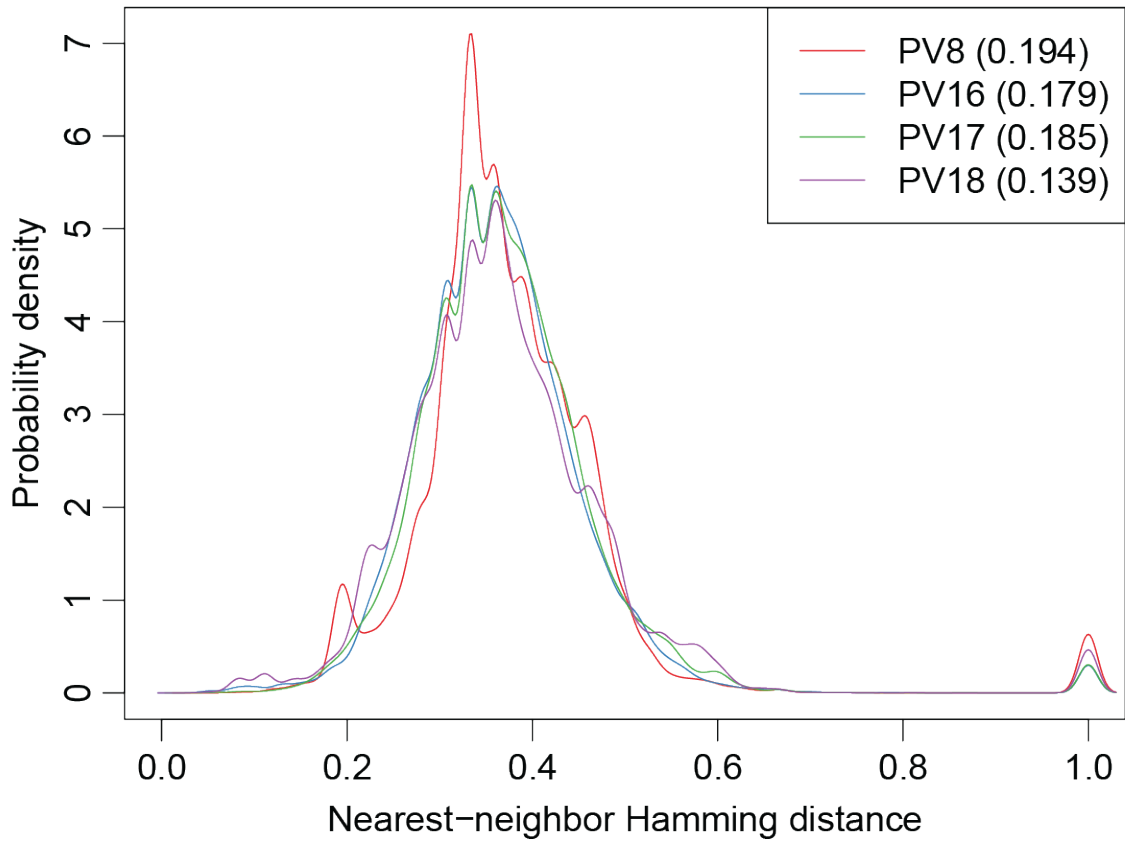
IgG1-IgA1 or IgG4-IgA2 parent-child pairs are cross-reactive to some antigen found in the normal flora or present on mucosal surfaces. Somewhat suggestive of this possibility is the fact that some anti-Dsg antibodies are cross-reactive to the rotavirus coat protein VP6 (401). A circulating anti-Dsg IgG1 or IgG4 may enter a mesenteric lymph node or Peyer's patches in the lamina propria, where it is induced to switch from IgG1 to IgA1 (and, potentially, subsequently IgA2) or from IgG4 to IgA2 by the presence of said antigen and a permissive cytokine milieu. The identity of said antigen(s) is beyond the scope of this study, though it could be determined by targeted sequencing of the mucosal IgA repertoire in PV and screening of anti-Dsg IgA clones for cross-reactivity against a library of mucosal antigens found in normal flora.

In conclusion, it appears that in PV, the majority of anti-Dsg IgG4 B cells show no evidence of arising from other subclasses, while anti-Dsg IgA1, IgG1, and IgA2 B cells show more extensive relationships among each other. These results also imply that sequential class switching is not a dominant pathway of IgG4 production in PV, despite its temporal relationship with IgG1 in the setting of chronic antigen stimulation. Further repertoire analysis and B cell lineage tracing in IgG4-centric conditions, like PV and other states of continuous antigen challenge, would be needed to shed light on this claim.

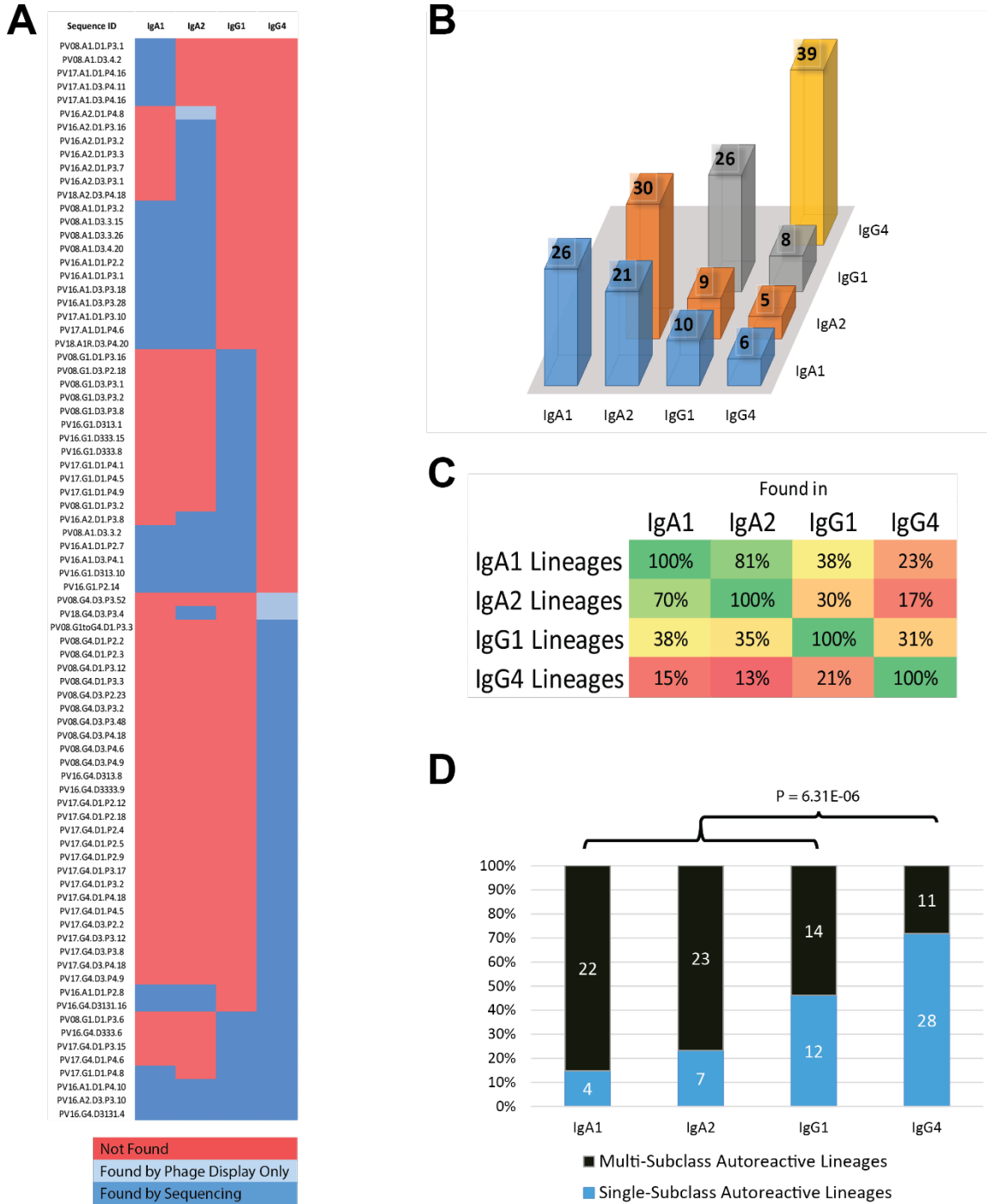
### 3.6 Figures and Tables



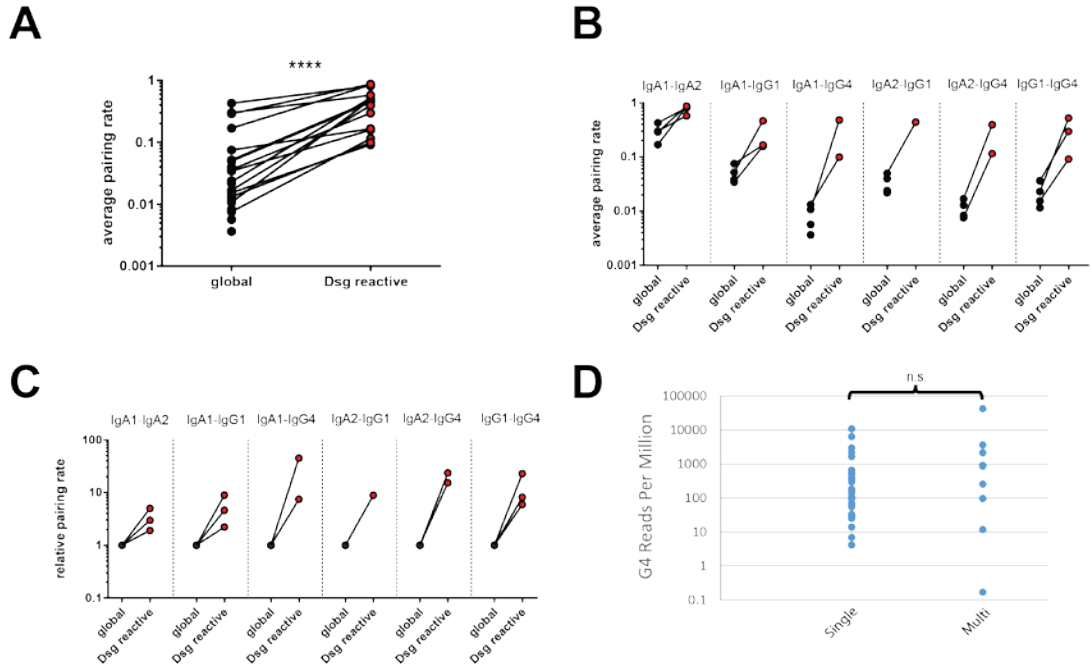
**Figure 3-1. Schematic of NGS library amplicon.** After an initial amplification with hinge-specific primers, subsequent amplification using constant-region-specific primers adds a library signature and Illumina adapters. After de-multiplexing, only reads with the matching isosignature were kept.



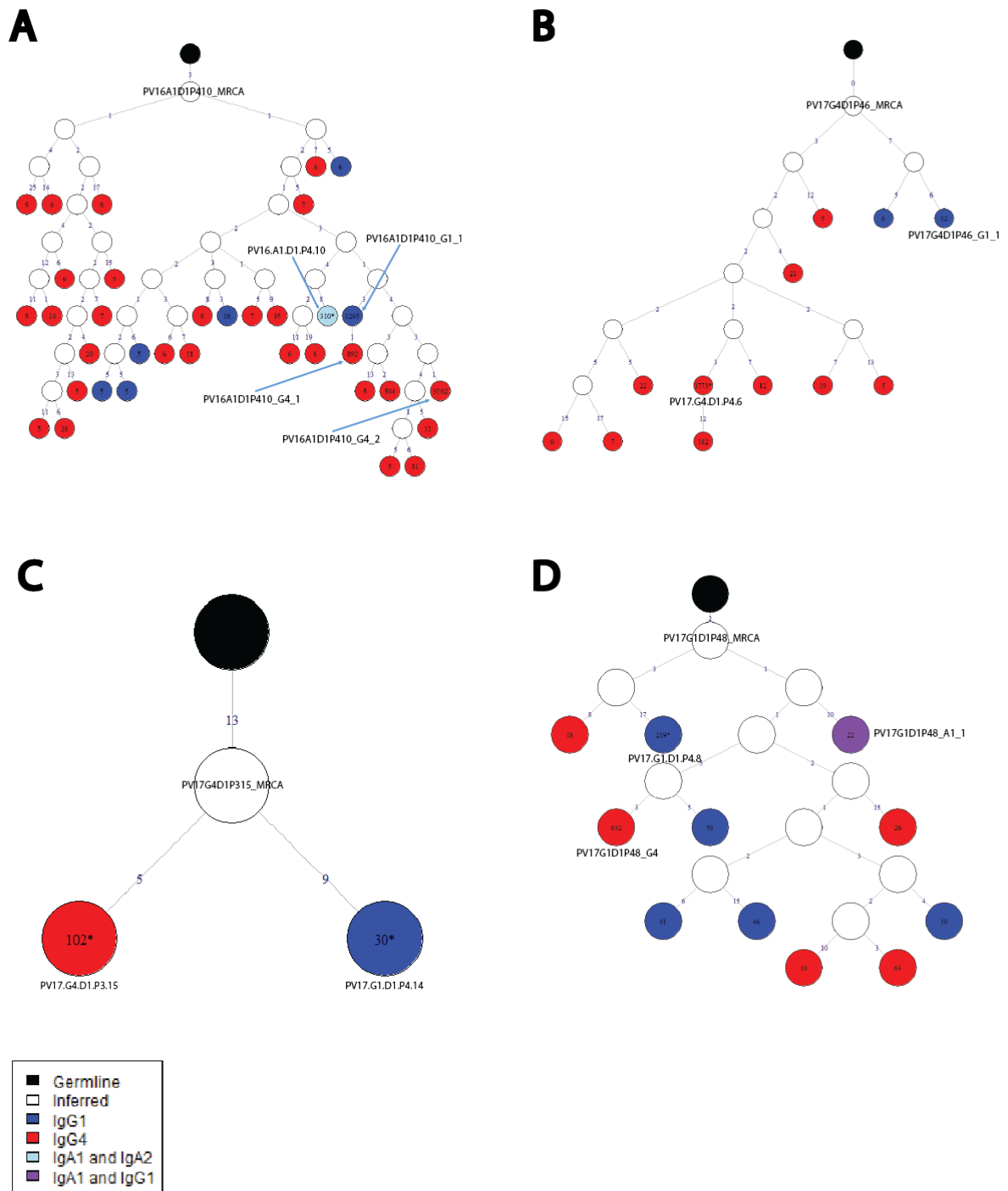
**Figure 3-2. Deriving CDR3 Bounds for Clonality.** After sorting all sequences by V gene, J gene, and CDR3 length, the nucleotide Hamming distance between CDR3 from one patient were compared to those from all other patients; since it is impossible for two people to have clonally related CDR3, this distribution was used to determine a 99% confidence interval a CDR3 dissimilarity cutoff. Values in parentheses were selected as cutoffs.



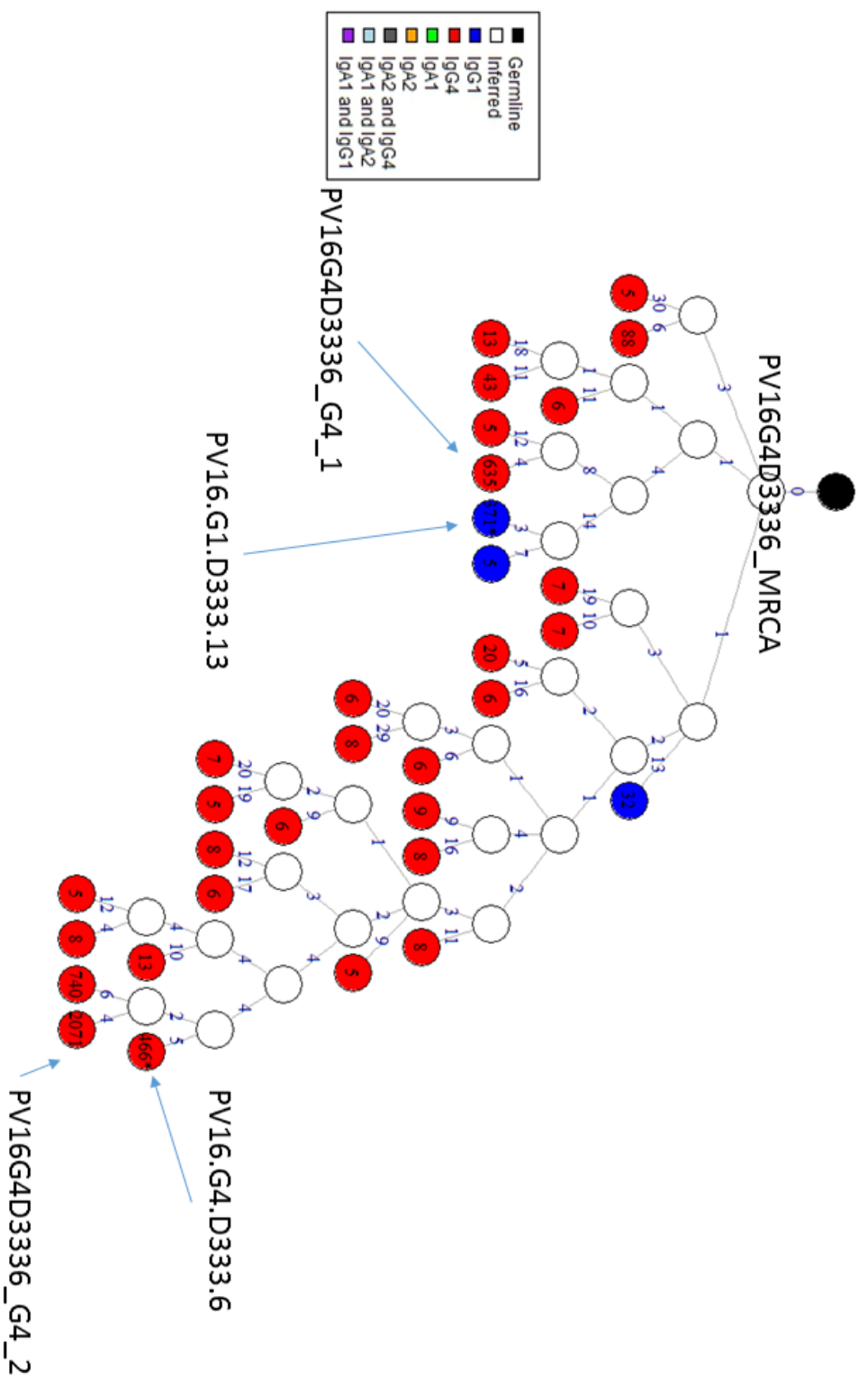
**Figure 3-3. Autoreactive Lineages Bridge Subclasses.** (A) Phage display identified 80 lineages that were then mapped back to NGS-derived lineages. Cross-subclass relationships are displayed in blue as shown. In total, we have found 26 IgA1-, 30 IgA2-, 26 IgG1-, and 39 IgG4-containing lineages. (B) Number of lineages containing every possible pair of subclasses. (C) Percentage of lineages containing a given subclass that have relatives in another. IgA1 and IgA2 lineages, in particular, show a great deal of overlap between them, while IgG1 and IgG4 do not. (D) Distribution of lineages among single-subclass and multi-subclass lineages within each subclass. IgG4 was compared to the sum of IgA1, IgA2, and IgG1 by 2 x 2 two-tailed Fisher exact test.



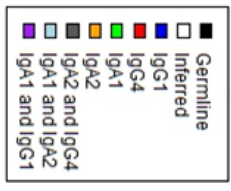
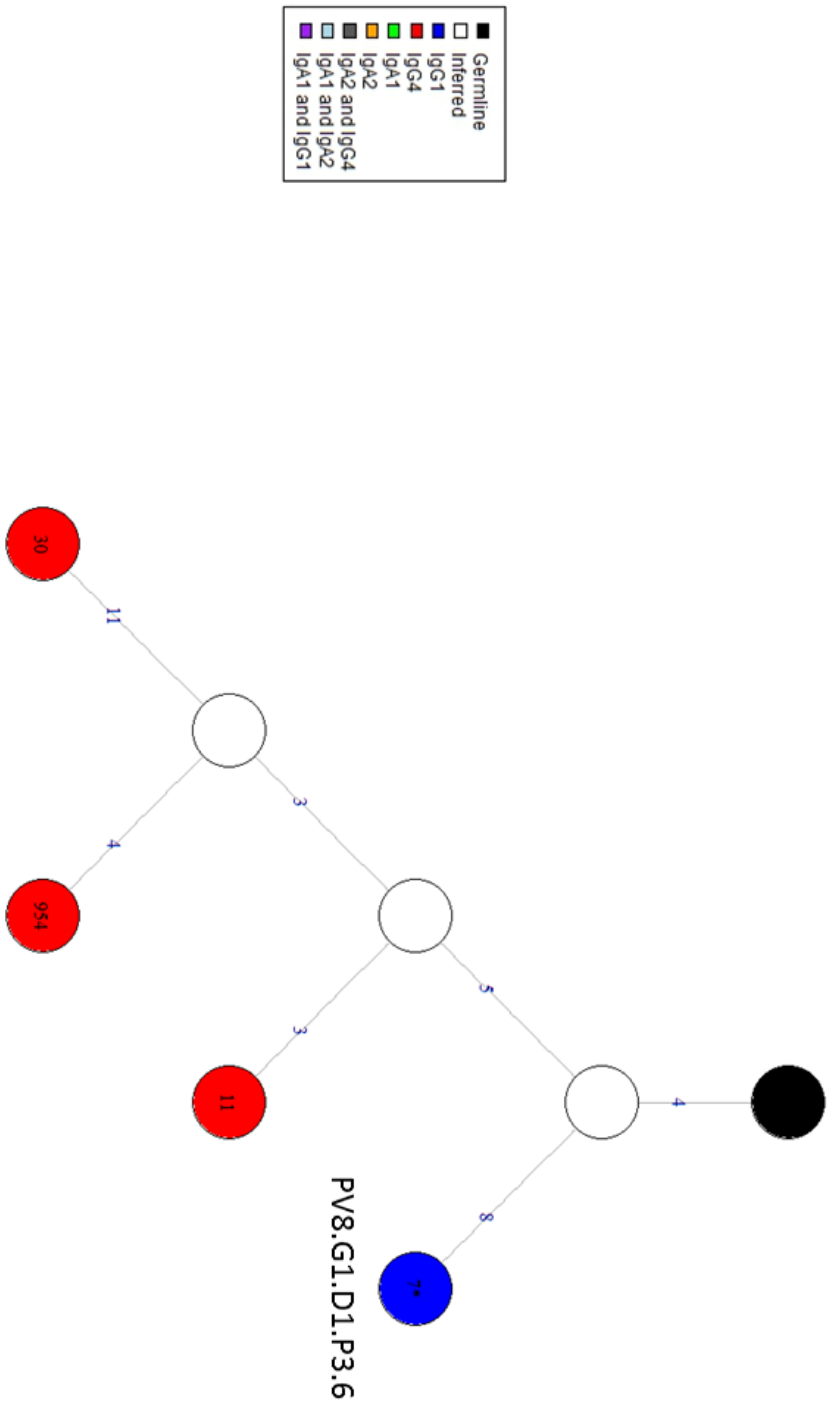
**Figure 3-4. Connectivity of the Dsg-Reactive and Global Repertoire.** (A) The average pairing rate for each subclass was plotted for both the global and Dsg-reactive repertoires. Each data point represents a pair of subclasses within a patient. The Dsg-reactive repertoire across 4 patients contains 14 (of a possible 24) pairs of subclasses. (B) Absolute and (C) relative pairing rates separated out by subclass pairs. Relative pairing rates are normalized to the global repertoire. (D) Read numbers for single- and multi-subclass IgG4-containing anti-Dsg lineages were compared to determine whether pairing rate is dependent upon number of reads mapping to a given lineage. Mann-Whitney test was not significant. PV18 was excluded because no PV18 IgG4 anti-Dsg clones mapped to IgG4-containing lineages derived from sequencing. n.s. - not significant.



**Figure 3-5. Selected Trees from IgG4-Containing Multi-Subclass Lineages.** All trees were generated using the maximum parsimony algorithm from PHYLIP after centroid clustering at 97% radius using USEARCH for each subclass individually; therefore, nodes representing more than one subclass arise from identical centroid sequences in the subclasses represented. Read numbers are written inside each node, with a \* indicating vertices mapping to a phage-display-derived sequence. Edge labels indicate number of mutations. Labelled vertices indicate sequences produced and tested by ELISA (see Figure 3-21). Lineages are (A) PV16.A1.D1.P4.10 (B) PV17.G4.D1.P4.6 (C) PV17.G4.D1.P3.15 (D) PV17.G1.D1.P4.8

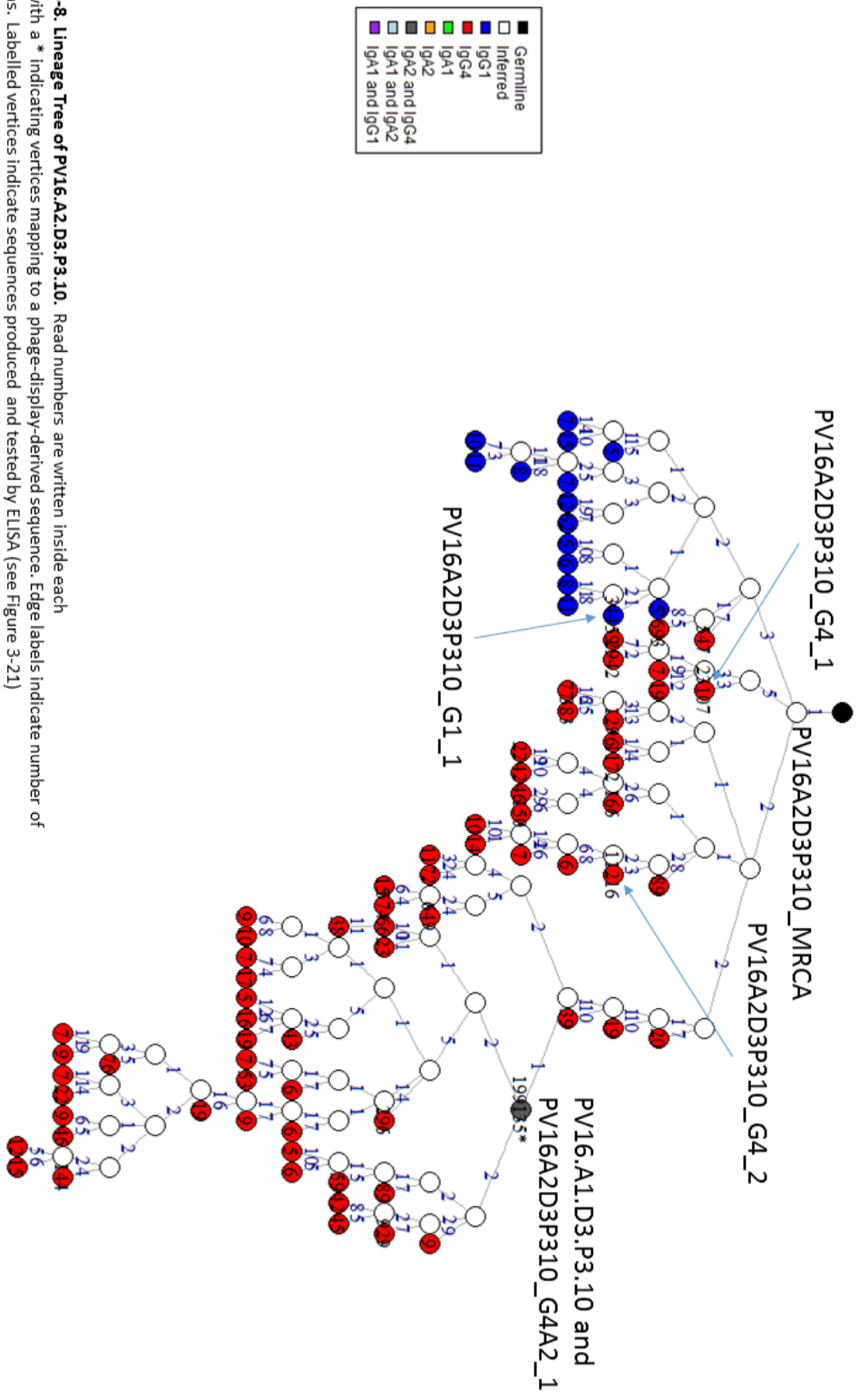


**Figure 3-6. Lineage Tree of PV16.G4.D333.6.** Read numbers are written inside each vertex, with a \* indicating vertices mapping to a phage-display-derived sequence. Edge labels indicate number of mutations. Labelled vertices indicate sequences produced and tested by ELISA (see Figure 3-21)

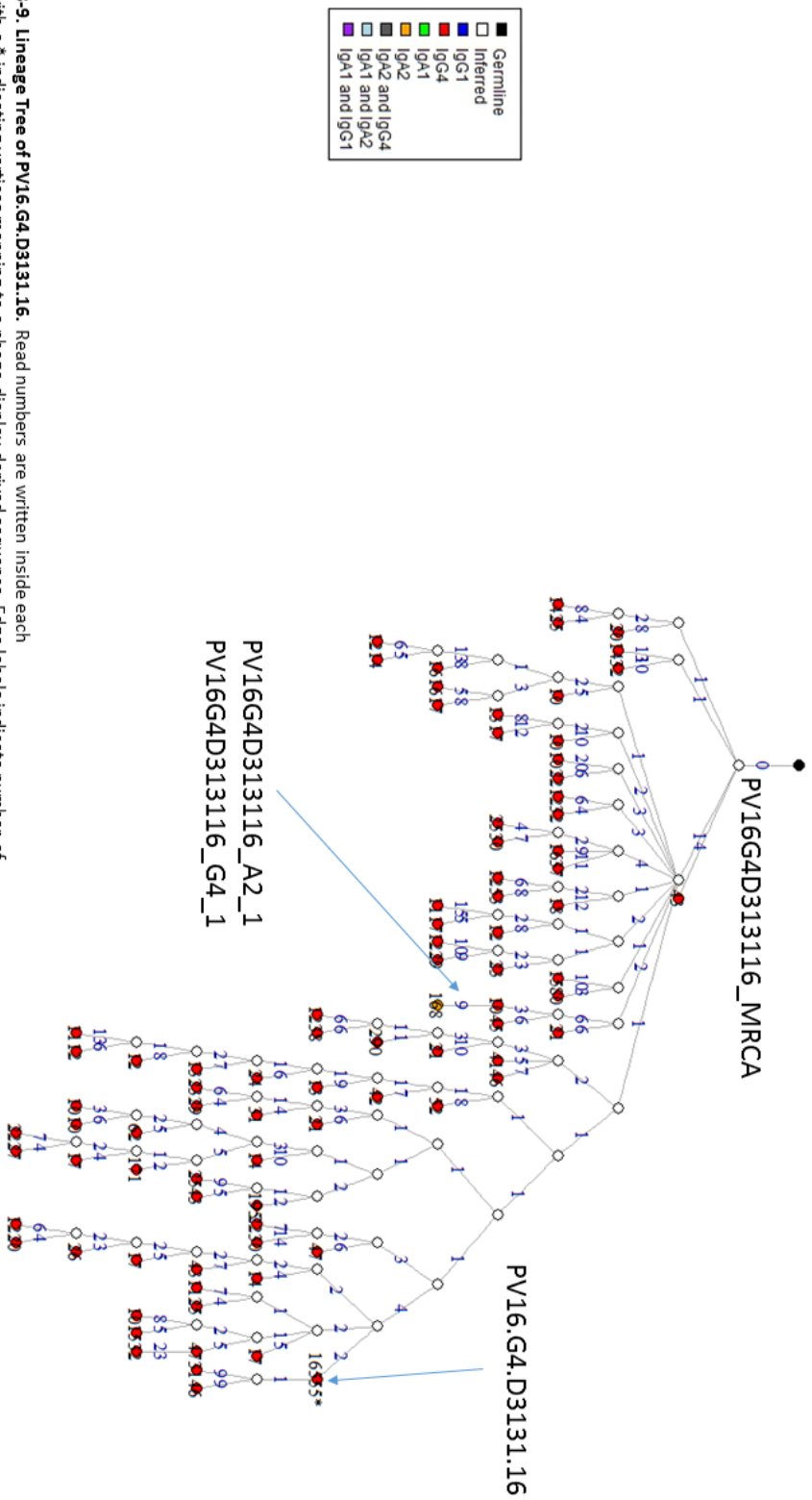


**Figure 3-7. Lineage Tree of PV8.G1.D1.P3.6.** Read numbers are written inside each vertex, with a \* indicating vertices mapping to a phage-display-derived sequence. Edge labels indicate number of mutations.

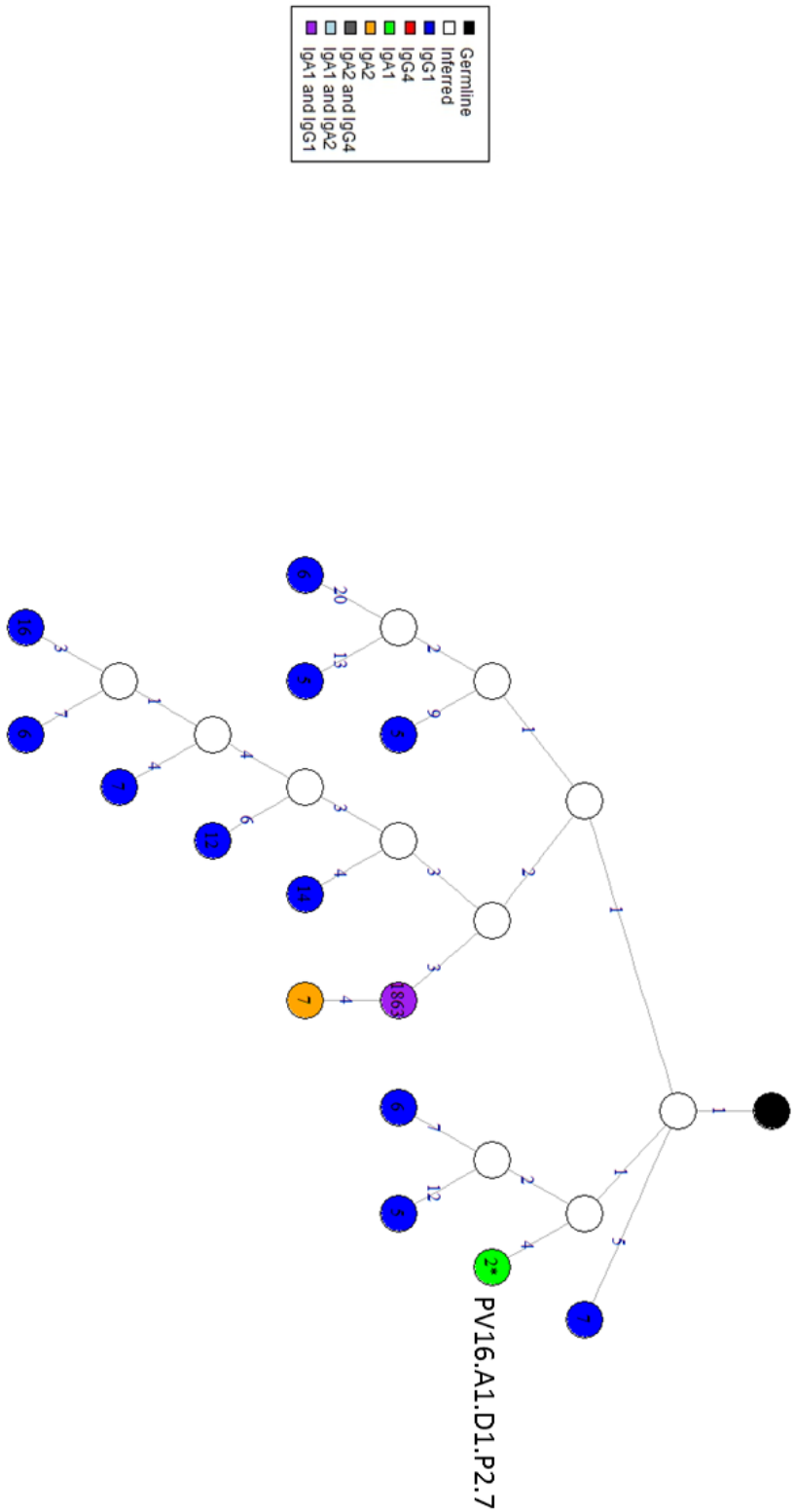




**Figure 3-8. Lineage Tree of PV16.A1.D3.P3.10 and PV16A2D3P310\_G4A2\_1.** Read numbers are written inside each vertex, with a \* indicating vertices mapping to a phage-display-derived sequence. Edge labels indicate number of mutations. Labelled vertices indicate sequences produced and tested by ELISA (see Figure 3-21)

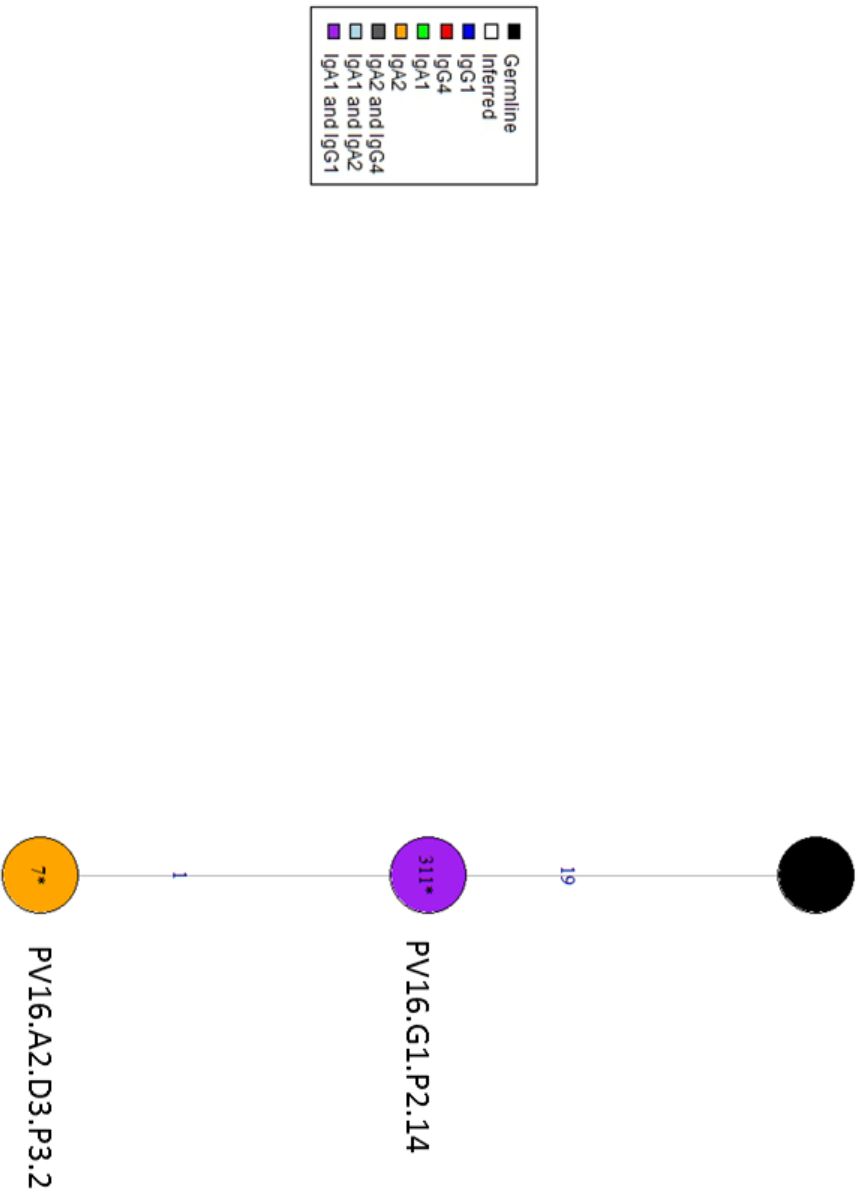


**Figure 3-9. Lineage Tree of PV16.G4.D3131.16.** Read numbers are written inside each vertex, with a \* indicating vertices mapping to a phage-display-derived sequence. Edge labels indicate number of mutations. Labelled vertices indicate sequences produced and tested by ELISA (see Figure 3-21). All nodes with readcount <10 were removed.

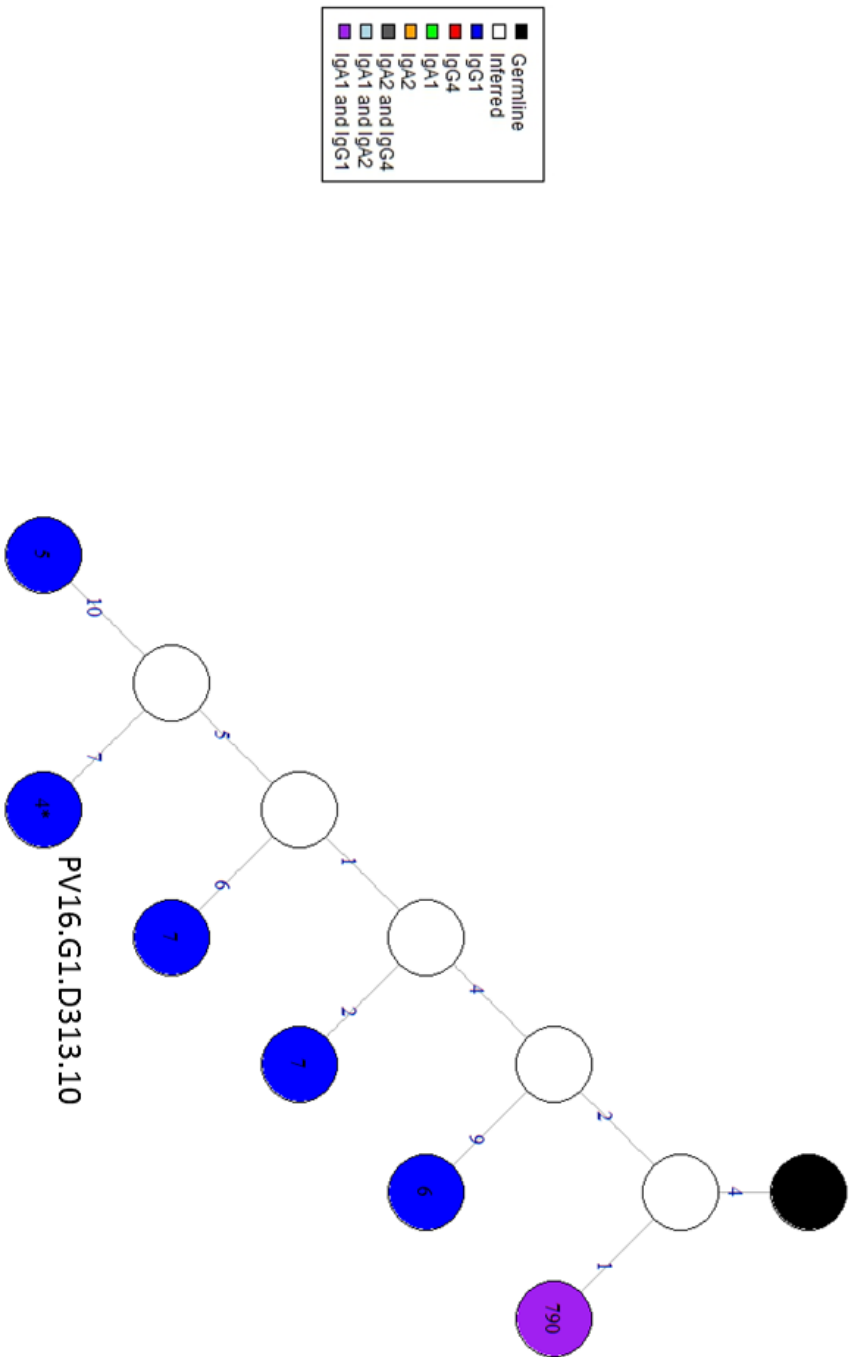


Black square	Germeline
White square	Inferred
Blue square	IgG1
Red square	IgG4
Green square	IgA1
Orange square	IgA2
Grey square	IgA2 and IgG4
Light blue square	IgA1 and IgA2
Purple square	IgA1 and IgG1

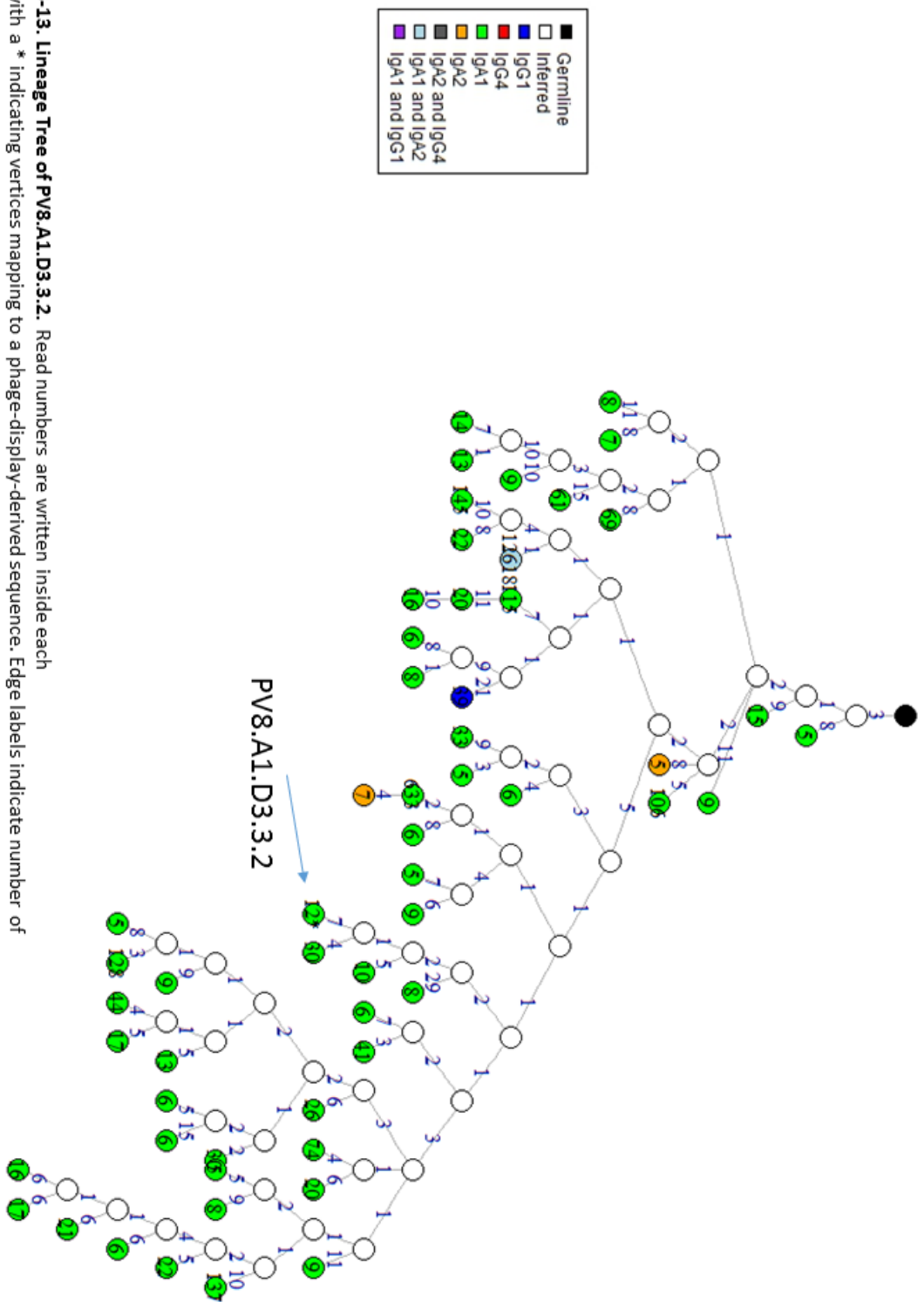
**Figure 3-10. Lineage Tree of PV16.A1.D1.P2.7.** Read numbers are written inside each vertex, with a \* indicating vertices mapping to a phage-display-derived sequence. Edge labels indicate number of mutations.



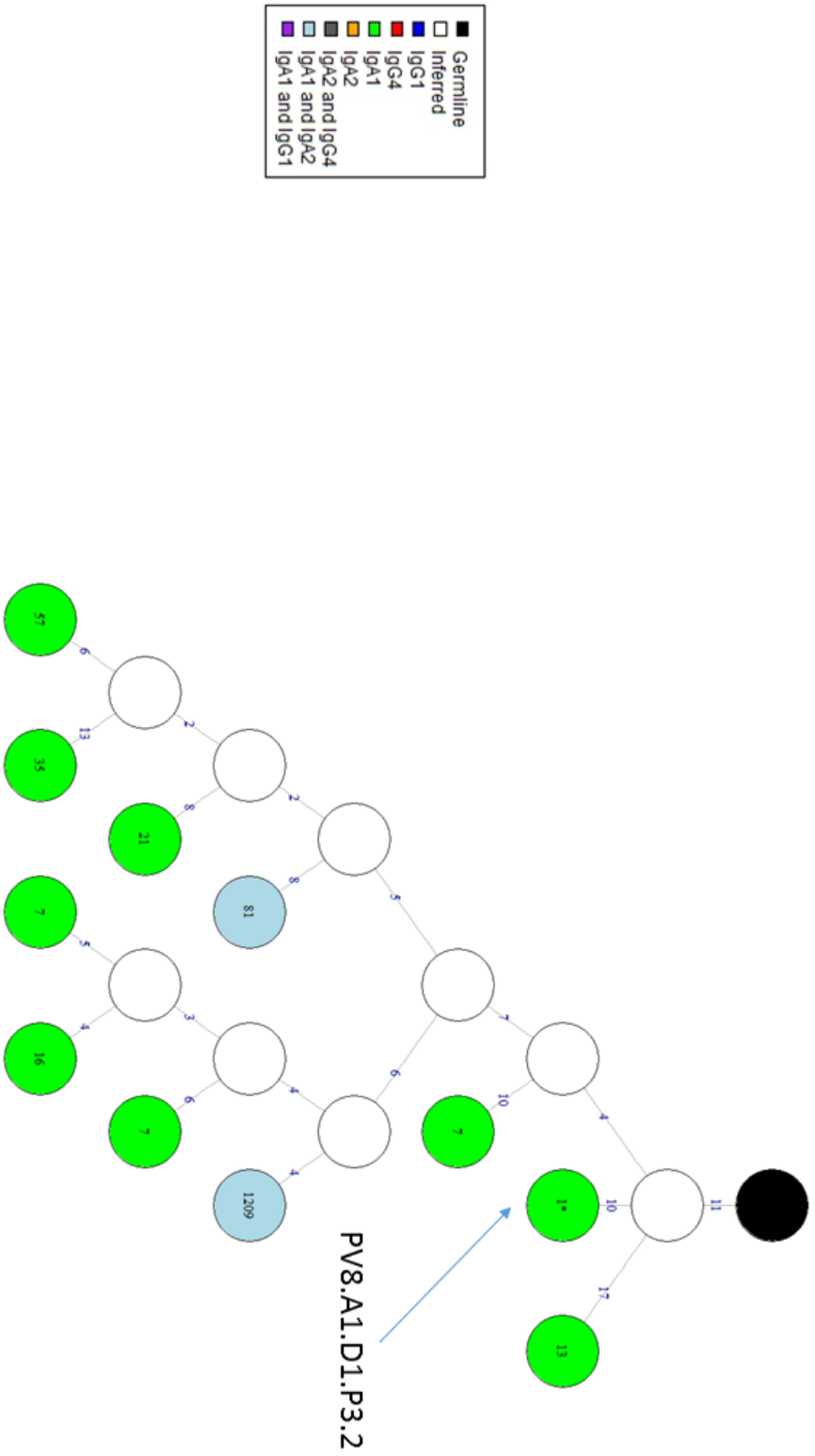
**Figure 3-11. Lineage Tree of PV16.G1.P2.14.** Read numbers are written inside each vertex, with a \* indicating vertices mapping to a phage-display-derived sequence. Edge labels indicate number of mutations.



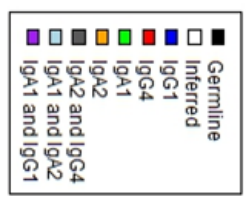
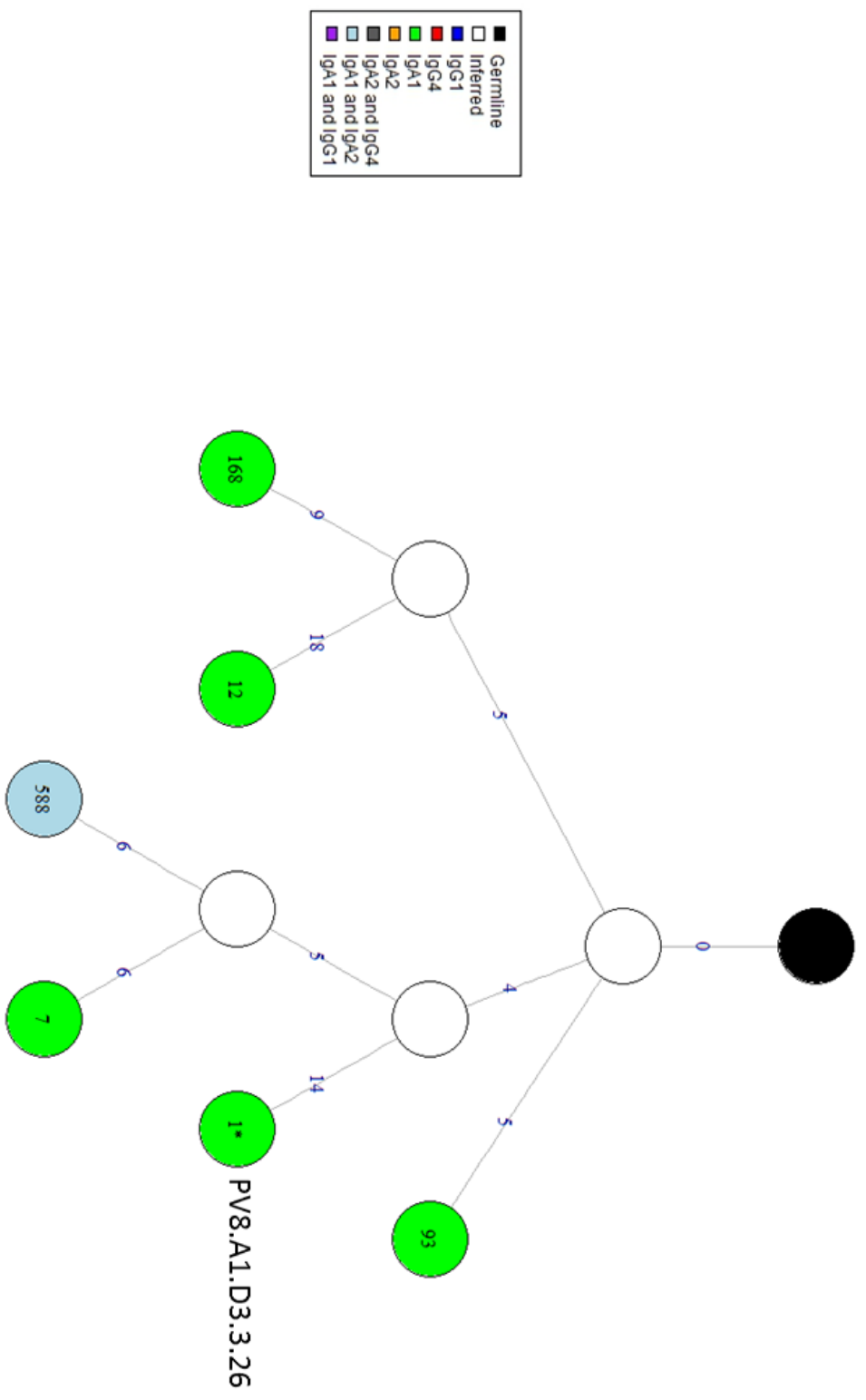
**Figure 3-12. Lineage Tree of PV16.G1.P313.10.** Read numbers are written inside each vertex, with a \* indicating vertices mapping to a phage-display-derived sequence. Edge labels indicate number of mutations.



**Figure 3-13. Lineage Tree of PV8.A1.D3.3.2.** Read numbers are written inside each vertex, with a \* indicating vertices mapping to a phage-display-derived sequence. Edge labels indicate number of mutations.

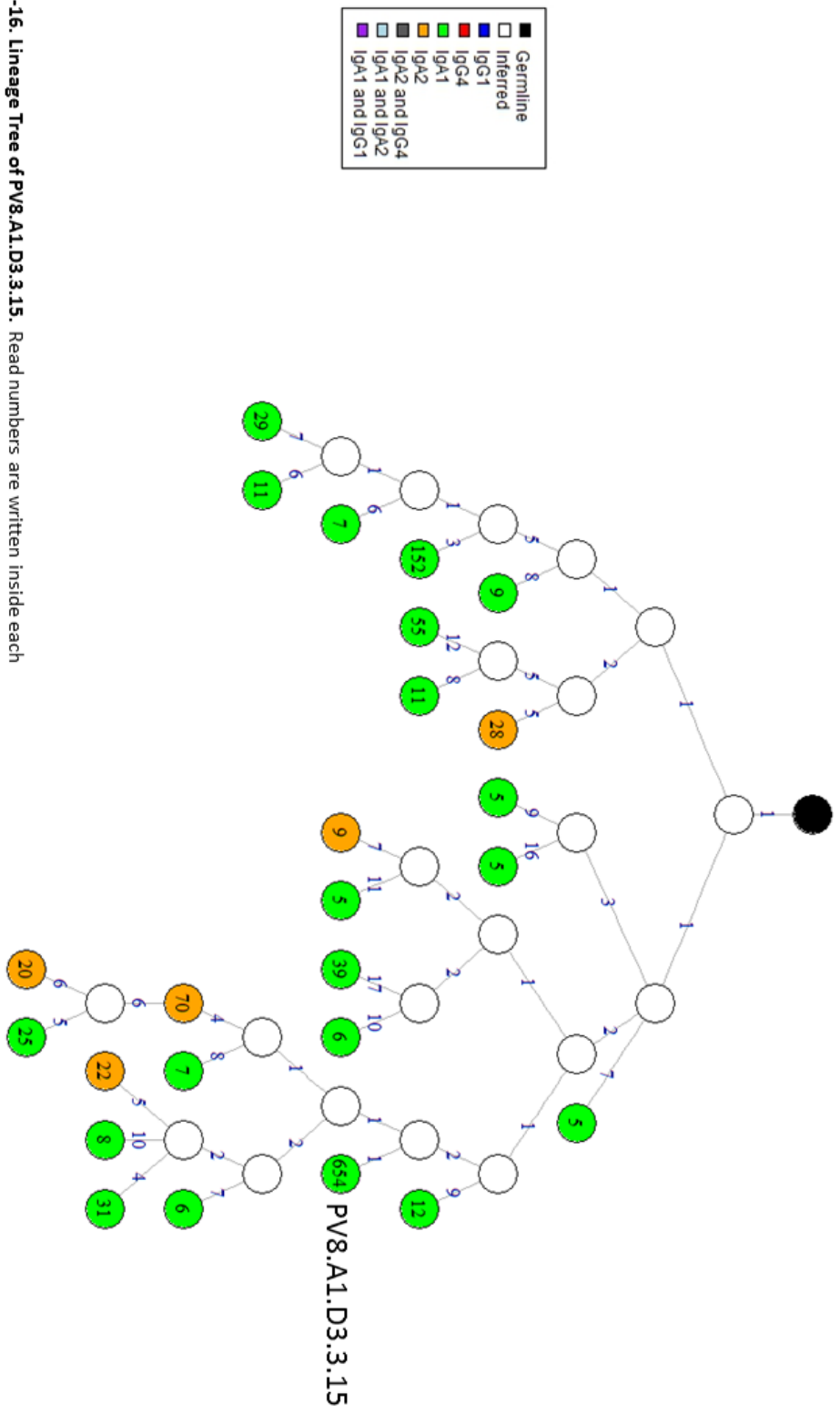


**Figure 3-14. Lineage Tree of PV8.A1.D1.P3.2.** Read numbers are written inside each vertex, with a \* indicating vertices mapping to a phage-display-derived sequence. Edge labels indicate number of mutations.

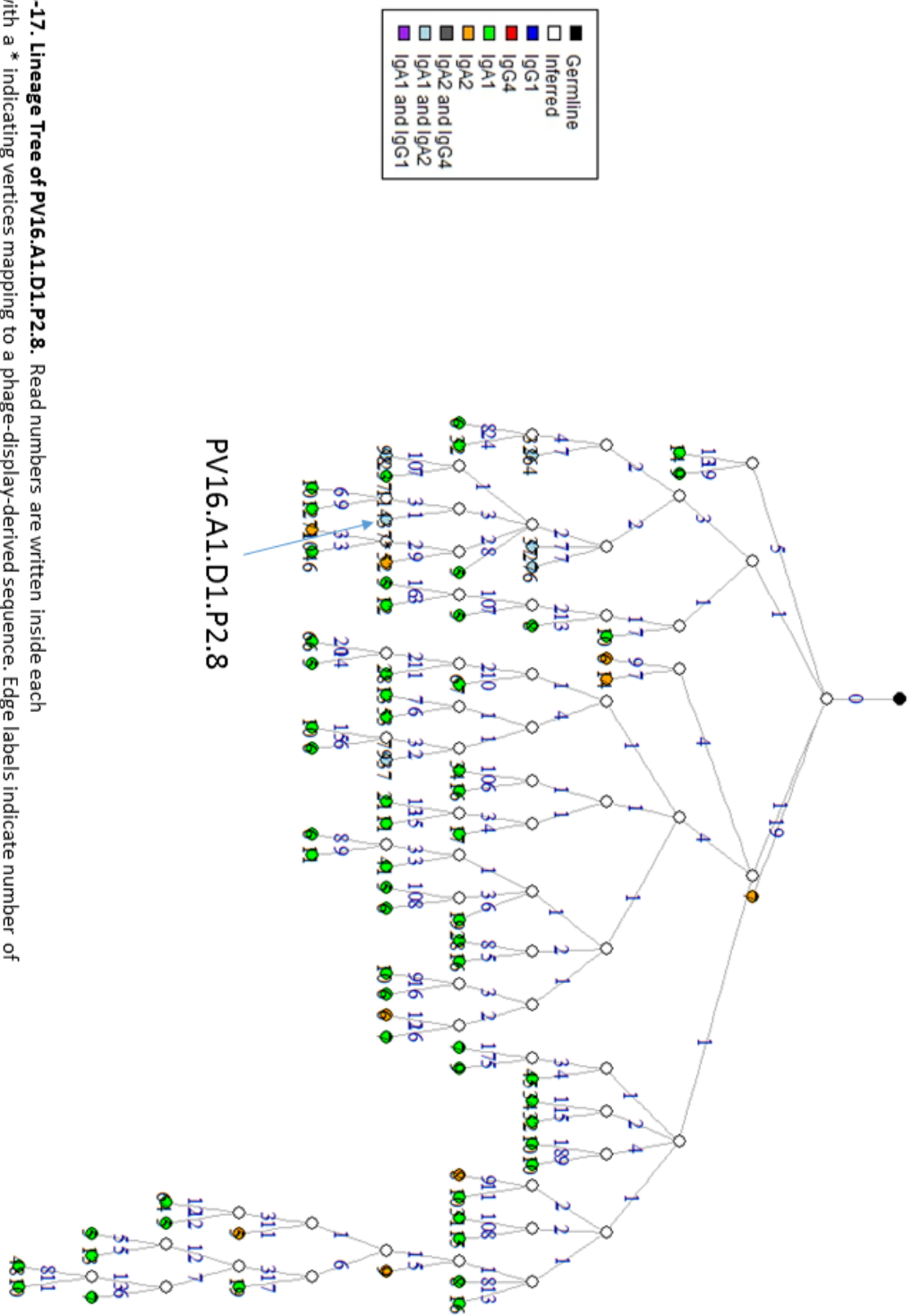


**Figure 3-15. Lineage Tree of PV8.A1.D3.3.26.** Read numbers are written inside each vertex, with a \* indicating vertices mapping to a phage-display-derived sequence. Edge labels indicate number of mutations.

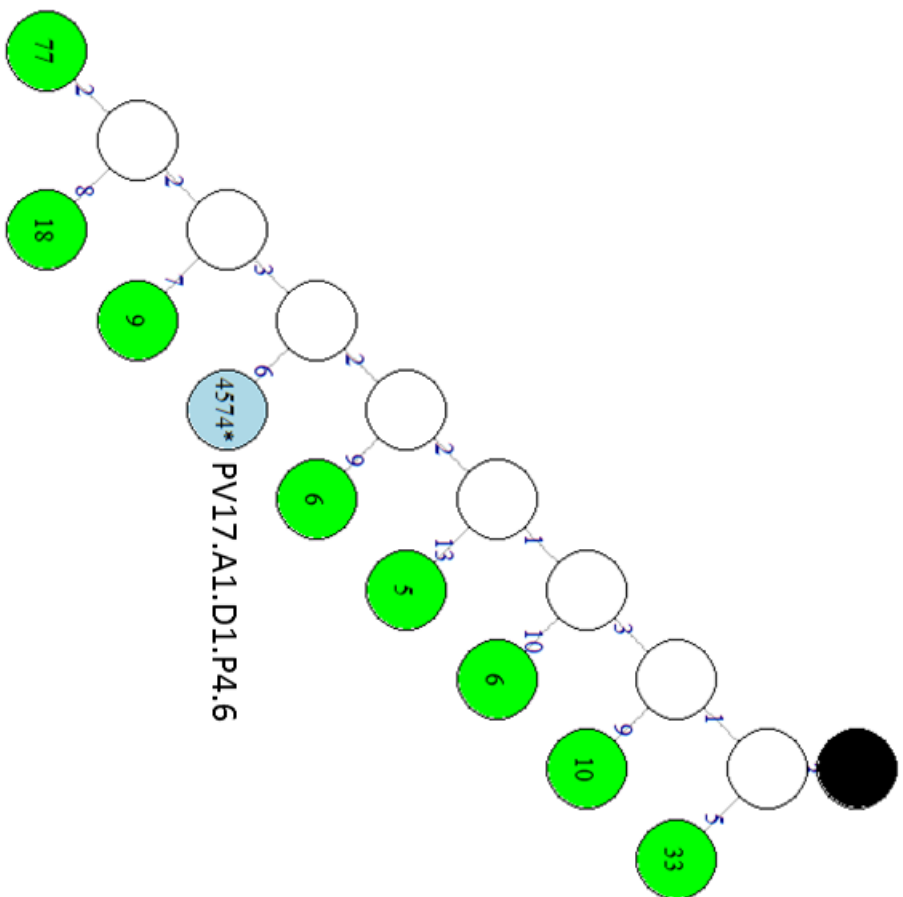
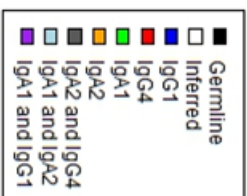




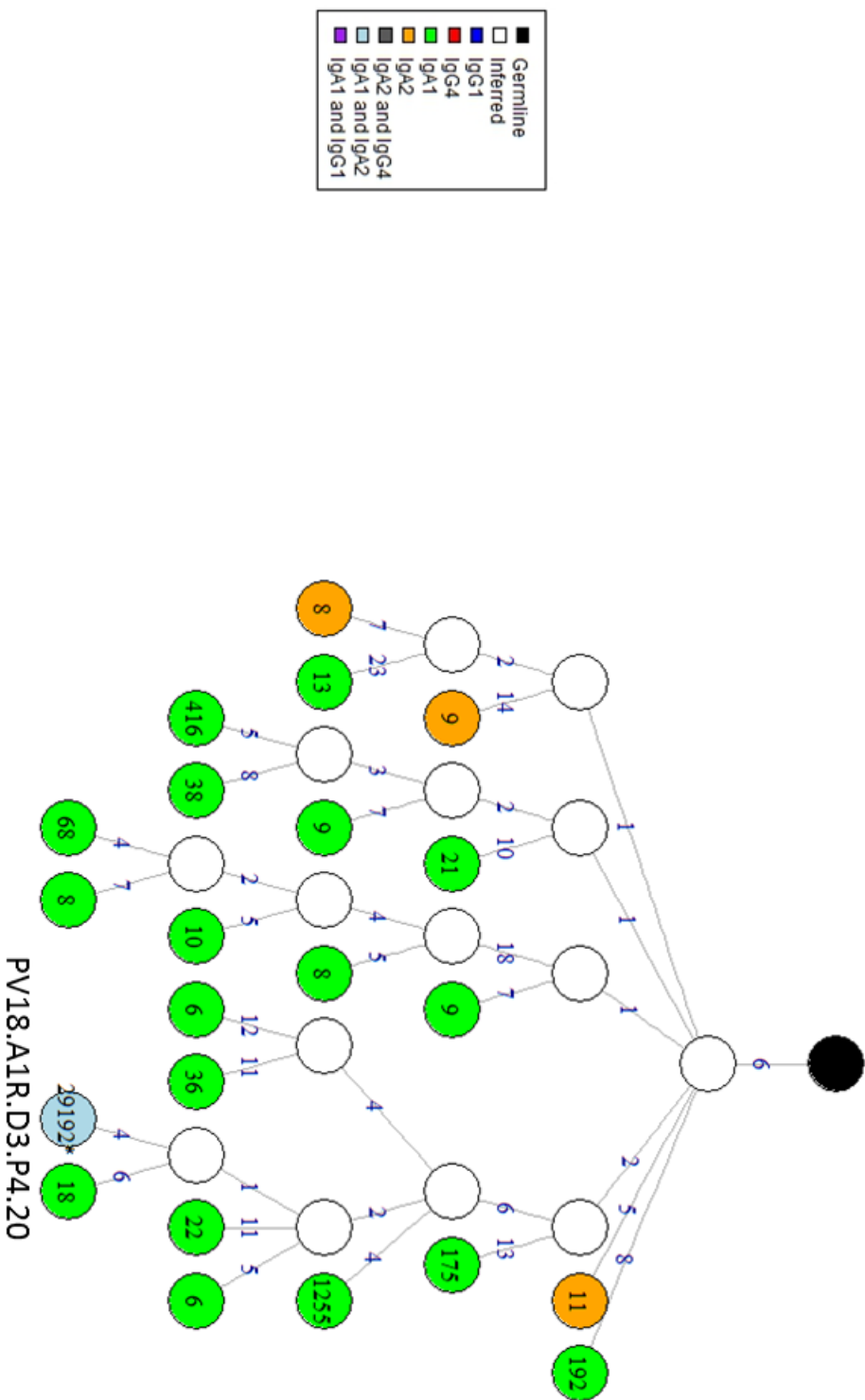
**Figure 3-16. Lineage Tree of PV8.A1.D3.3.15.** Read numbers are written inside each vertex, with a \* indicating vertices mapping to a phage-display-derived sequence. Edge labels indicate number of mutations.



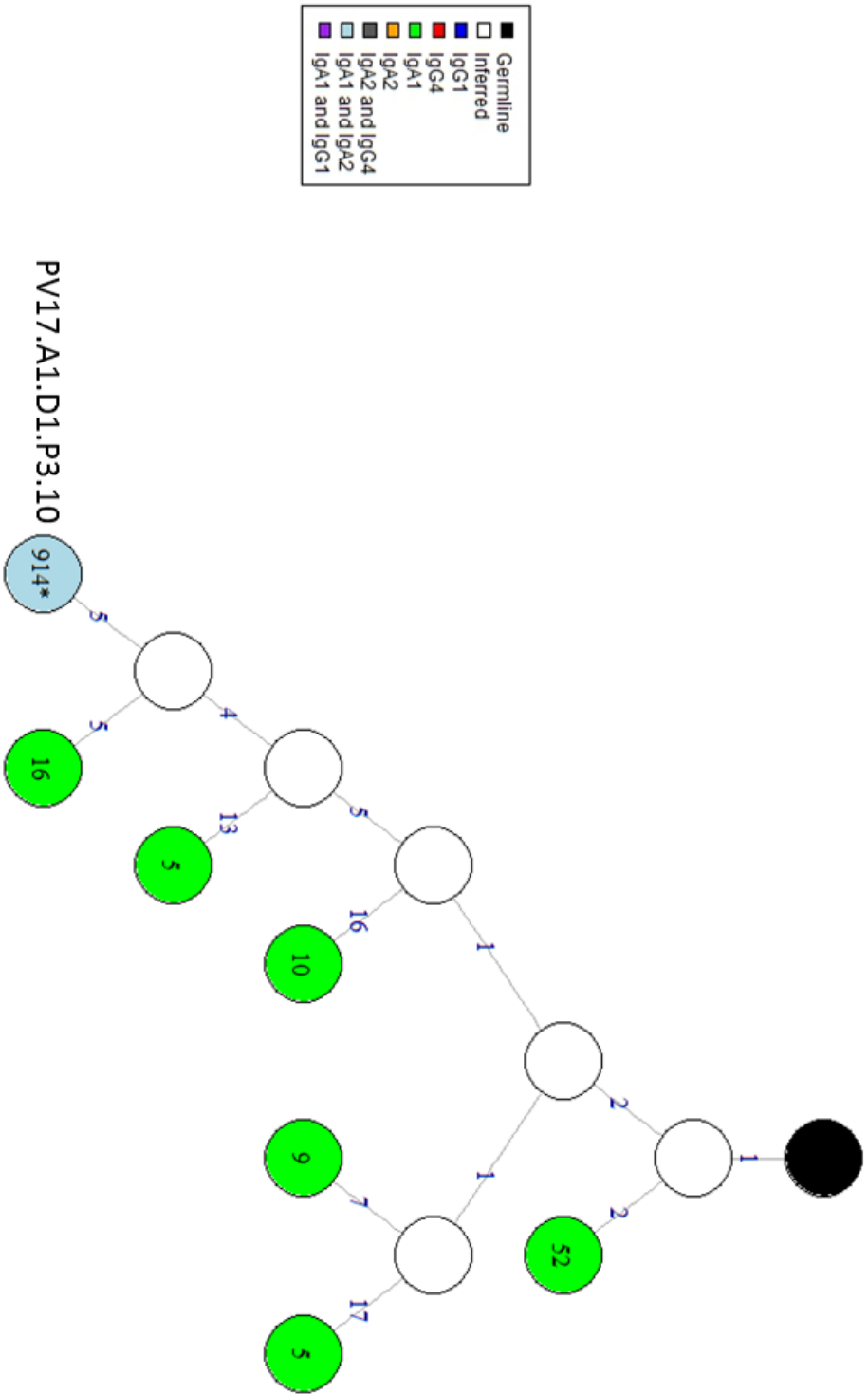
**Figure 3-17. Lineage Tree of PV16.A1.D1.P2.8.** Read numbers are written inside each vertex, with a \* indicating vertices mapping to a phage-display-derived sequence. Edge labels indicate number of mutations.



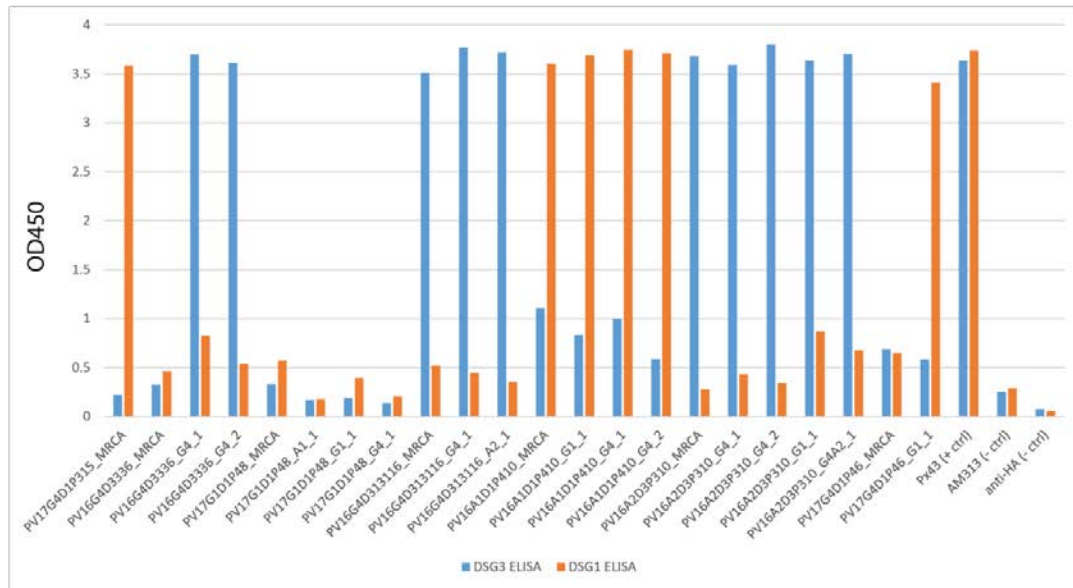
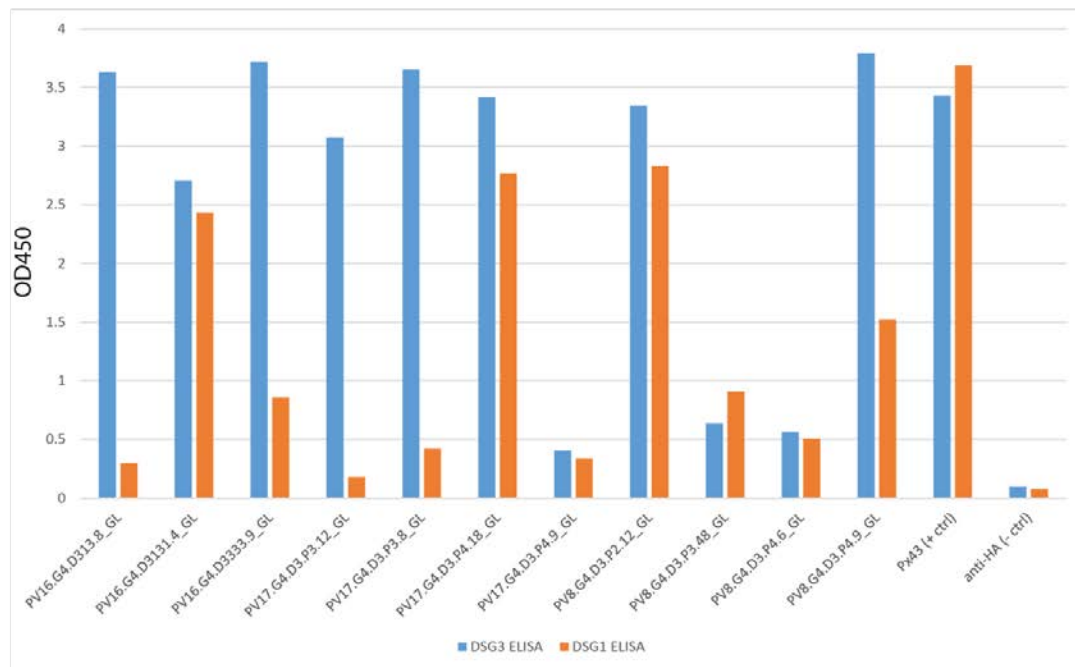
**Figure 3-18. Lineage Tree of PV17.A1.D1.P4.6.** Read numbers are written inside each vertex, with a \* indicating vertices mapping to a phage-display-derived sequence. Edge labels indicate number of mutations.



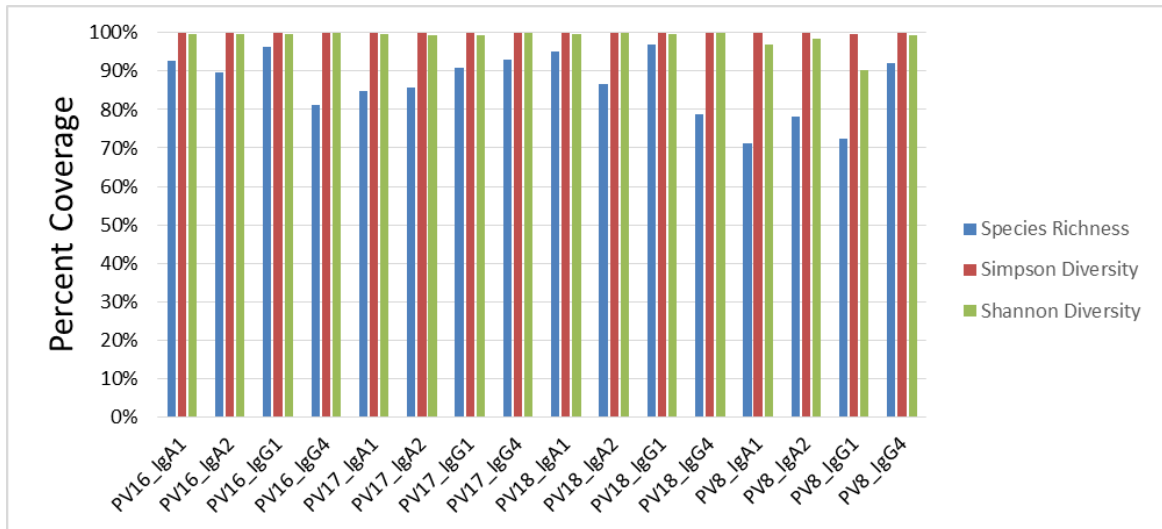
**Figure 3-19. Lineage Tree of PV18.A1R.D3.P4.20.** Read numbers are written inside each vertex, with a \* indicating vertices mapping to a phage-display-derived sequence. Edge labels indicate number of mutations.



**Figure 3-20. Lineage Tree of PV17.A1.D1.P3.10.** Read numbers are written inside each vertex, with a \* indicating vertices mapping to a phage-display-derived sequence. Edge labels indicate number of mutations.

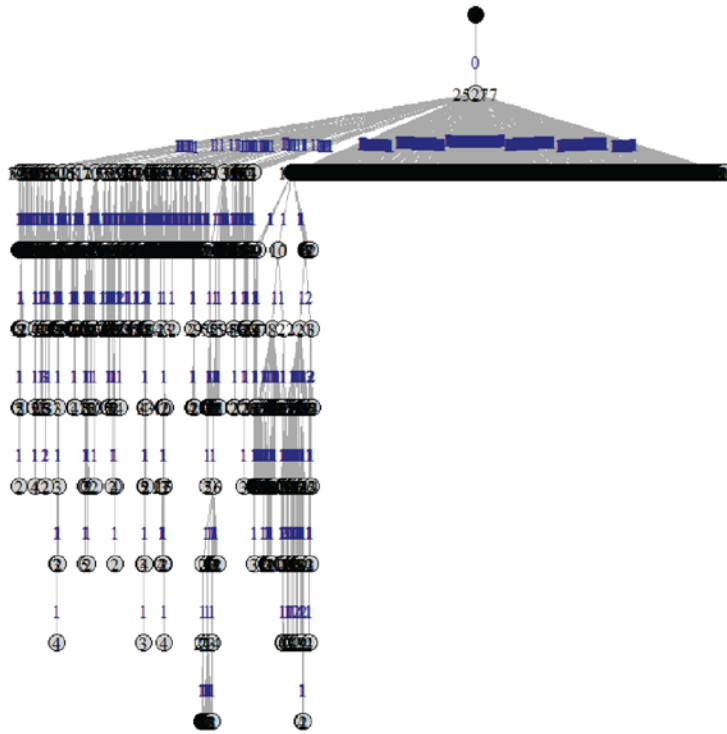
**A****B**

**Figure 3-21. ELISA reactivity of cloned lineage intermediates and germline-reverted clones.** (A) Relatives from IgG4-containing multi-subclass lineages (See Figures 3, Supplemental Figure 1 and 2 for identities) were tested at 10 ug/mL. (B) Germline-reverted IgG4 clones were tested at 10 ug/mL. GL = Germline, MRCA = Most Recent Common Ancestor.



**Figure 3-22. Quantification of Sampling Coverage by Asymptotic Diversity Estimation.**

Coverage is calculated by dividing the given diversity measure against the asymptotic estimation of that measure, i.e. the value of that measure had we sampled each patient-subclass pair infinitely. In this way, we can determine the percentage of the total of any given diversity measure that we have captured. Results indicate that we have captured between 71-97% of the species richness, meaning 71-97% of all lineages. The percentage of overall species richness captured is lower than the percentage of Shannon or Simpson diversity captured, because Shannon and Simpson diversity weight common species more than rare species.

**A****B**

**Figure 3-23. Error Correction For Lineage Tree Production.** A monoclonal template was amplified using the same protocols as the subclass-specific sequencing libraries. (A) Before any error correction, there are 1443 non-redundant sequences. (B) After USEARCH clustering with 97% radius and elimination of all sequences with below 5 reads, over 99.99% of reads are identical to the original template.



Patient	Clinical Description	Dsg1 ELISA	Dsg3 ELISA	IIF	Comment
PV8	Severe Mucocutaneous PV	IgA1+, IgA2+, IgG1+/-, IgG4+	IgA1+, IgA2+, IgG1+, IgG4+	IgG+, IgA questionable	Patient on Treatment (Mycophenolate/Prednisone)
PV16	Mucocutaneous PV	IgA1+, IgA2+, pan-IgG+	IgA1+, IgA2+, pan-IgG+	IgG+, IgA1+, IgA2+	New Onset Disease, No Systemic Therapy
PV17	Mucocutaneous PV	pan-IgG+ (Index 50)	pan-IgG+ (Index 178)	[Not Available]	7 Month History of Disease, No Systemic Therapy
PV18	Mucosal PV	pan-IgG- (Index 5)	pan-IgG+ (Index 187)	[Not Available]	6 year history of disease treated with repeated steroid tapers. Off of systemic therapy at time of blood draw

**Table 3-1. Patient Characteristics.** All ELISA and IIF results shown are reported from day of blood draw.

Original Sequence ID	Original Subclass	Subclasses Represented in Lineages	Direct Switches	Figure
PV16.A1.D1.P4.10	A1	A1, A2, G1, G4	A1-A2, G1-G4	3-5a
PV17.G4.D1.P4.6	G4	G1, G4	None	3-5b
PV17.G4.D1.P3.15	G4	G1, G4	None	3-5c
PV17.G1.D1.P4.8	G1	A1, G1, G4	G1-A1	3-5d
PV16.G4.D333.6	G4	G1, G4	None	3-6
PV08.G1.D1.P3.6	G1	G1, G4	None	3-7
PV16.A2.D3.P3.10	A2	A2, G1, G4	G4-A2	3-8
PV16.G4.D3131.16	G4	A2, G4	G4-A2	3-9
PV16.A1.D1.P2.7	A1	A1, A2, G1	G1-A1, A1-A2, G1-A2	3-10
PV16.G1.P2.14	G1	A1, A2, G1	G1-A1, A1-A2, G1-A2	3-11
PV16.G1.D313.10	G1	A1, G1	G1-A1	3-12
PV08.A1.D1.P3.2	A1	A1, A2	A1-A2	3-13
PV08.A1.D3.3.2	A1	A1, A2, G1	A1-A2	3-14
PV08.A1.D3.3.26	A1	A1, A2	A1-A2	3-15
PV08.A1.D3.3.15	A1	A1, A2	None	3-16
PV16.A1.D1.P2.8	A1	A1, A2	A1-A2	3-17
PV17.A1.D1.P4.6	A1	A1, A2	A1-A2	3-18
PV18.A1R.D3.P4.20	A1	A1, A2	A1-A2	3-19
PV17.A1.D1.P3.10	A1	A1, A2	A1-A2	3-20

**Table 3-2. Index of Lineage Trees.** Each of the 19 trees analyzed in the results section are listed here. Switches were considered direct iff the representative sequences from the two subclasses were identical, or they were connected by a single edge.



PV8 Lineage #	SEQID	SUBCLASS	HC V-GENE	HC J-GENE	AA HCDBS JUNCTION	HC CDR3 LENGTH	LC V-GENE	LC J-GENE	AA LCDBS JUNCTION	LC JUNCTION LENGTH
1	PV08.A1.D1.P3.1	IgA1	IGHV3-23	IGH4	ATHRWSD	7	IGLV6-57	IGJ2	GSYDSTNVV	9
2	PV08.A1.D1.P3.2	IgA1	IGHV1-18	IGH4	AKDDPNRGLVDS	12	IGHV4-1	IGK4	QCYTGGTPP	8
3	PV08.A1.D3.1.15	IgA1	IGHV3-15	IGH5	ARHGKAGVSD	12	IGLV6-57	IGJ3	GSYDTMDDV	9
4	PV08.A1.D3.1.2	IgA1	IGHV3-15	IGH5	ATDOGRH	7	IGKV1-11	IGK4	QQRVNAJLFT	10
5	PV08.A1.D3.1.26	IgA1	IGHV2-5	IGH4	AHTYYYSGSSFPYDFD	19	IGLV6-57	IGJ3	GSYDSSHHQV	10
6	PV08.A1.D3.1.4.2	IgA1	IGHV2-5	IGH4	ARTTNSNFY	11	IGLV6-57	IGJ2	GSYDSSNVL	9
7	PV08.A1.D3.1.20	IgA1	IGHV3-30	IGH4	ARGALVNHGFWSSPPGEY	18	IGLV1-51	IGJ3	GTWDSLSLQGV	11
8	PV08.G1.D1.P3.16	IgG1	IGHV3-48	IGH6	ARQKWVNLGTYTTHMDV	18	IGLV7-43	IGJ3	LYVYGGAGPWP	11
9	PV08.G1.D1.P3.2	IgG1	IGHV1-69	IGH6	ARGHSSLSFYVYVDS	15	IGLV6-57	IGJ2	GSYNSSHHV	10
10	PV08.G1.D1.P3.1	IgG1	IGHV1-48	IGH3	ARQVLDLSSGYSAAHAFDI	20	IGKV1-39	IGK5	QDQYSLSST	10
11	PV08.G1.D3.P2.18	IgG1	IGHV2-5	IGH3	VHTPWVVGDFDI	13	IGLV1-40	IGJ2	GSYDLSRAVV	11
12	PV08.G1.D3.P3.1	IgG1	IGHV5-51	IGH3	ARVGLGARAGVDFM	14	IGKV1-44	IGJ3	AAWDSLSNGAV	11
13	PV08.G1.D3.P3.2	IgG1	IGHV4-4	IGH4	ARGTTAFVQVGHY	15	IGLV2-30	IGK1	MGTWHPVT	9
14	PV08.G1.D3.P3.8	IgG1	IGHV4-34	IGH6	ARISPYEILTGSMDV	16	IGLV2-14	IGJ2	SSYSSSPV	10
15	PV08.G1.D3.P3.1	IgG1	IGHV7-4-1	IGH6	ARIGFTGLGFRHVMOL	18	IGLV2-14	IGJ2	SSYTSSTVYV	12
16	PV08.G4.D1.P2.1	IgG4	IGHV7-4-1	IGH4	ARGGLWHSLEYSYVGS	18	IGHV1-5	IGK2	GDYNSPPT	9
17	PV08.G4.D1.P2.1	IgG4	IGHV1-18	IGH4	ARREHHHRSGLGRHYMS	19	IGKV2-14	IGJ1	SSYTSPL	9
18	PV08.G4.D1.P2.4	IgG4	IGHV7-4-1	IGH6	ARGGFTGLGFRHVMOL	18	IGLV6-57	IGJ2	GSYDSSHHV	10
19	PV08.G4.D1.P3.12	IgG4	IGHV3-73	IGH6	ARNGKWTPTQYVYVMDV	18	IGHV1-5	IGK2	GDYNSYGVV	9
20	PV08.G4.D1.P3.1	IgG4	IGHV1-18	IGH4	AREEAVAGSSLSHWLTSYDFD	25	IGLV1-44	IGJ2	TAWDLSLQGV	12
21	PV08.G4.D3.P2.12	IgG4	IGHV2-5	IGH6	ALSHRYTDKWSGTYVYVMDV	20	IGHV1-5	IGK1	GDYSYTPW	9
22	PV08.G4.D3.P2.19	IgG4	IGHV4-4	IGH6	ARPPHVDLSDGTYVYVMDV	21	IGLV2-14	IGJ7	SSYTSSTLV	10
23	PV08.G4.D3.P2.23	IgG4	IGHV4-41	IGH4	ARVRLWAGGYSYVDFD	20	IGHV8-61	IGJ3	VLWVSSDQV	10
24	PV08.G4.D3.P2.1	IgG4	IGHV2-5	IGH6	ALSHRYTQSDTYVYVMDV	20	IGHV8-61	IGJ3	LYVLDSDW	10
25	PV08.G4.D3.P3.40	IgG4	IGHV2-5	IGH6	ALSHRYTQSSGTYVYVMDV	20	IGHV1-5	IGK3	QDYSYPLT	9
26	PV08.G4.D3.P3.48	IgG4	IGHV3-9	IGH4	AKDAGGVESDPLDY	15	IGKV1-39	IGK4	QDYSYPLT	9
27	PV08.G4.D3.P3.52	IgG4	IGHV4-61	IGH6	ARGARWTKGNFGEFHYYMDEI	21	IGLV1-44	IGJ2	GTWDSLSLQGV	11
28	PV08.G4.D3.P4.18	IgG4	IGHV4-4	IGH6	ARHPVLYDFSSGDTVYVMDV	21	IGKV3-21	IGJ2	GTWDSRAHFL	11
29	PV08.G4.D3.P4.6	IgG4	IGHV1-69	IGH3	ARKEISTNDAFDM	13	IGLV2-14	IGJ2	SSYTSNTPPV	11
30	PV08.G4.D3.P4.9	IgG4	IGHV2-5	IGH6	ARSGVGLDTPH	13	IGLV1-40	IGJ7	GSYDSSSLQGV	12
PV16 Lineage #	SEQID	SUBCLASS	HC V-GENE	HC J-GENE	AA HCDBS JUNCTION	HC CDR3 LENGTH	LC V-GENE	LC J-GENE	AA LCDBS JUNCTION	LC JUNCTION LENGTH
1	PV16.A1.D1.P2.2	IgA1	IGHV1-46	IGH4	ARAFVAGTDD	12	IGLV6-57	IGJ2	GSYDSTNVV	10
2	PV16.A1.D1.P2.7	IgA1	IGHV1-46	IGH6	AREHWYDSDGTYVYVMDV	19	IGKV3-15	IGK5	QDYNWPLLT	10
3	PV16.A1.D1.P2.8	IgA1	IGHV2-5	IGH4	AHRLWNFPQYDFD	14	IGLV4-69	IGJ2	GTWGLGRV	9
4	PV16.A1.D1.P3.1	IgA1	IGHV4-31	IGH4	ARVSHYHNNAYNYDFD	17	IGLV1-51	IGJ3	GTWDSLSLQVW	11
5	PV16.A1.D1.P4.10	IgA1	IGHV3-30-3	IGH5	ARSGFPRNLDGSGRGNWFDP	21	IGKV3-15	IGK1	QDYNWPLPT	10
6	PV16.A1.D3.P1.18	IgA1	IGHV3-33	IGH4	ARFVLEWDFSDYDFD	15	IGLV2-14	IGJ2	SSYTSSTLV	9
7	PV16.A1.D3.P2.28	IgA1	IGHV1-2	IGH5	ARGVGTGNPWFDP	14	IGLV1-47	IGJ3	AAWDSLSLQGV	11
8	PV16.A1.D3.P4.1	IgA1	IGHV3-48	IGH4	ARVYVDSAAHWG	14	IGKV1-44	IGJ3	AATWDLGRV	11
9	PV16.A2.D1.P3.16	IgA2	IGHV1-46	IGH6	ARGRLGLTSSDYNESYGLDV	22	IGLV6-57	IGJ3	GSYDSSNVL	9
10	PV16.A2.D1.P3.2	IgA2	IGHV3-30	IGH6	VKDRWDWVYKGTGSMDV	17	IGLV2-28	IGK2	NQGLQTPY	9
11	PV16.A2.D1.P3.3	IgA2	IGHV4-31	IGH4	ARAGGQDLSYDFD	13	IGLV1-44	IGJ3	ATWDDRLQGV	11
12	PV16.A2.D1.P3.4	IgA2	IGHV2-5	IGH6	AREHWYDSDGTYVYVMDV	19	IGKV3-15	IGK2	QDYNWPLPT	10
13	PV16.A2.D1.P3.6	IgA2	IGHV3-30-3	IGH5	ARSGFPRNLDGSGRGNWFDP	21	IGKV3-15	IGK1	QDYNWPLPT	10
14	PV16.A2.D1.P3.7	IgA2	IGHV3-33	IGH6	ARGVYSGVYVGMDDV	16	IGLV1-44	IGJ2	AAWDSLSLQV	11
15	PV16.A2.D1.P3.8	IgA2	IGHV2-5	IGH3	ARVWDDTFD	11	IGHV1-5	IGK1	QDYSYPLT	9
16	PV16.A2.D1.P4.8	IgA2	IGHV4-4	IGH4	VKSICGGGAFES	12	IGKV2-28	IGK4	MCALQPLT	9
17	PV16.A2.D3.P3.1	IgA2	IGHV2-5	IGH6	ATGYTBMDEI	10	IGLV2-14	IGJ2	SSYTSSTLVV	11
18	PV16.A2.D3.P3.10	IgA2	IGHV3-34	IGH6	ARSPNSABHYVYVMDV	17	IGKV1-17	IGK4	LQHNVPPT	9
19	PV16.A2.D3.P3.13	IgA2	IGHV2-5	IGH3	ARVWDDTFD	11	IGKV1-27	IGK4	QDYNWPLPT	9
20	PV16.A2.D3.P3.15	IgA2	IGHV2-5	IGH3	ARVWDDTFD	11	IGKV1-39	IGK2	QDYSYPLT	9
21	PV16.A2.D3.P3.2	IgA2	IGHV3-66	IGH4	ARDLVGSYDERDYSYGLDY	21	IGKV1-12	IGK4	QDANSPLT	9
22	PV16.A2.D3.P3.3	IgA2	IGHV2-5	IGH3	ARVWDDTFD	11	IGLV6-57	IGJ3	GSYDSSNVL	9
23	PV16.G1.D3.13.1	IgG1	IGHV3-73	IGH6	ARHGAGDFD	11	IGKV1-12	IGK1	QDANSFPT	9
24	PV16.G1.D3.13.10	IgG1	IGHV3-73	IGH6	TSRGLNLSYVGMDDV	17	IGKV1-39	IGK1	QDYSYPLT	9
25	PV16.G1.D3.13.15	IgG1	IGHV1-69	IGH3	ARALRPTHLPWYDFD	18	IGKV3-20	IGK5	QDYSYPLT	9
26	PV16.G1.D3.13.15	IgG1	IGHV1-18	IGH4	ARDYDGGRCDFY	14	IGHV1-5	IGK1	QDYNWPLPT	9
27	PV16.G1.D3.13.8	IgG1	IGHV1-69	IGH3	ARNSYTTGADF	12	IGKV3-21	IGJ3	QVWDSSSDHPV	11
28	PV16.G1.P2.14	IgG1	IGHV3-66	IGH4	ARDLVGSYDERDYSYGLDY	21	IGKV1-39	IGK4	QDYSYPLT	9
29	PV16.G4.D3.13.8	IgG4	IGHV3-9	IGH4	ARVWGEYTSYDFD	16	IGLV2-8	IGJ3	SSYAGSNVL	10
30	PV16.G4.D3.13.16	IgG4	IGHV3-33	IGH3	ARDVNAFD	9	IGKV3-21	IGJ2	QVWDSSSDHPV	11
31	PV16.G4.D3.13.14	IgG4	IGHV3-30	IGH6	ARELPLPRATYCTNGATYSYGLDV	27	IGKV3-20	IGK3	QDYSYPLT	10
32	PV16.G4.D3.13.6	IgG4	IGHV1-69	IGH3	ARMLPQVCSRTYVYVMDV	18	IGHV3-20	IGK4	QDYSYPLT	9
33	PV16.G4.D3.13.9	IgG4	IGHV2-5	IGH6	ARELPLPRATYCTNGATYSYGLDV	27	IGKV3-21	IGJ2	QVWDSSSDHPV	11
PV17 Lineage #	SEQID	SUBCLASS	HC V-GENE	HC J-GENE	AA HCDBS JUNCTION	HC CDR3 LENGTH	LC V-GENE	LC J-GENE	AA LCDBS JUNCTION	LC JUNCTION LENGTH
1	PV17.A1.D1.P3.10	IgA1	IGHV2-5	IGH5	ALFRGGDAGWDFD	14	IGLV2-14	IGJ2	SSYTSSTLVV	11
2	PV17.A1.D1.P4.16	IgA1	IGHV1-2	IGH4	ARDQVYDSTGFY	15	IGKV1-39	IGK5	QDYSYMPIT	10
3	PV17.A1.D1.P4.6	IgA1	IGHV2-26	IGH2	ARHGADDEYSSYRPLSGVDFD	14	IGLV1-44	IGJ2	AAWDSLSLQGV	11
4	PV17.A1.D3.P4.11	IgA1	IGHV4-4	IGH4	VDRDRTSRTGDS	24	IGLV1-51	IGJ2	GTWDSLSLQV	11
5	PV17.A1.D3.P4.16	IgA1	IGHV3-23	IGH6	AKGGYSYVGMDDV	13	IGKV1-39	IGK5	QDYSYPLT	10
6	PV17.G1.D1.P4.9	IgG1	IGHV7-5	IGH5	ALFRGGDAGWDFD	14	IGHV1-5	IGK1	ATWDSLTPV	11
7	PV17.G1.D1.P4.1	IgG1	IGHV3-30-3	IGH5	ARARVIRGALYDTH	15	IGLV3-21	IGJ2	QVWDSRHHV	9
8	PV17.G1.D1.P4.14	IgG1	IGHV1-3	IGH6	ARGGTFRGGVRRHNSGMDI	19	IGHV3-20	IGK4	QDYSYPLT	9
9	PV17.G1.D1.P4.5	IgG1	IGHV1-2	IGH4	A2TSCAGLFSY	12	IGLV1-51	IGJ2	GTWDSLSLQV	11
10	PV17.G1.D1.P4.8	IgG1	IGHV3-21	IGH6	ARDELTEDENSSYQTRKYVYVMDV	26	IGKV3-20	IGK2	QDYSYPLT	11
11	PV17.G1.D1.P4.9	IgG1	IGHV4-59	IGH4	ARVWNLTAFL	11	IGLV3-21	IGJ2	QVWDSRHHV	12
12	PV17.G4.D1.P2.12	IgG4	IGHV1-69	IGH3	ASRPGLTYSYVMDV	17	IGKV3-20	IGK2	QDYNWPLPT	9
13	PV17.G4.D1.P2.18	IgG4	IGHV3-39	IGH6	ARVWVTFGGALQVYVMDV	20	IGKV2-28	IGK2	MKGKQTPPT	9
14	PV17.G4.D1.P2.4	IgG4	IGHV1-46	IGH6	ARAFTYVAGVGMDDV	15	IGLV3-21	IGJ7	QVWDSRHHV	11
15	PV17.G4.D1.P2.5	IgG4	IGHV1-46	IGH6	ARVALPVGMDV	12	IGLV3-21	IGJ2	GAWDSNVV	8
16	PV17.G4.D1.P2.9	IgG4	IGHV1-46	IGH6	ARGFLQPLDY	12	IGLV3-21	IGK3	QVWDSRHHV	12
17	PV17.G4.D1.P3.15	IgG4	IGHV1-3	IGH6	ARGGTFGGVRRHNSGMDV	19	IGKV3-20	IGK2	QDYSYPLT	10
18	PV17.G4.D1.P3.17	IgG4	IGHV1-2	IGH4	VRSFRGLDFY	11	IGKV1-33	IGK1	QDYNWPLPT	9
19	PV17.G4.D1.P3.2	IgG4	IGHV3-21	IGH6	ARDVYSDYDLTGSSDQTYVYVMDV	26	IGLV3-25	IGJ7	QDYSYPLT	11
20	PV17.G4.D1.P4.1	IgG4	IGHV4-39	IGH6	ALFRGGDAGWDFD	14	IGHV1-5	IGK1	ATWDSLTPV	11
21	PV17.G4.D1.P4.18	IgG4	IGHV1-69	IGH6	A2HKGSSYQVYVGMDDV	16	IGLV6-57	IGJ2	QDYSYPLT	9
22	PV17.G4.D1.P4.5	IgG4	IGHV1-3	IGH3	ARGGVFKSNLWVGGDFD	19	IGKV1-39	IGK3	QDYSYPLT	9
23	PV17.G4.D1.P4.6	IgG4	IGHV1-8	IGH6	ARGHRYDLSFLPTIYVGMDDV	24	IGLV3-21	IGJ2	QVWDSRHHV	11
24	PV17.G4.D3.P2.2	IgG4	IGHV4-4	IGH6	ARDAFFDLTDAVGHSGMDV	22	IGHV4-1	IGK3	QDYSYPLT	9
25	PV17.G4.D3.P3.12	IgG4	IGHV3-49	IGH3	ARVPGGLWLEGDIAFD	17	IGLV3-21	IGJ2	QVWDSRHHV	11
26	PV17.G4.D3.P3.7	IgG4	IGHV4-4	IGH6	ARVAFRDLTDRHPTGMDV	22	IGHV4-1	IGK1	QDYSYPLT	9
27	PV17.G4.D3.P3.8	IgG4	IGHV3-23	IGH6	ALVARYDLSYVGMDDV	21	IGKV3-20	IGK2	QDYSYPLT	10
28	PV17.G4.D3.P4.18	IgG4	IGHV1-46	IGH3	AREAYRHLCTDFD	13	IGLV2-8	IGJ2	SSYAGSNVL	10
29	PV17.G4.D3.P4.9	IgG4	IGHV2-26	IGH6	ARIGTYDILGDSYVGMDDV	21	IGLV1-44	IGJ2	AAWDSLSLQV	11
PV18 Lineage #	SEQID	SUBCLASS	HC V-GENE	HC J-GENE	AA HCDBS JUNCTION	HC CDR3 LENGTH	LC V-GENE	LC J-GENE	AA LCDBS JUNCTION	LC JUNCTION LENGTH
1	PV18.A1R.D3.P4.20	IgA1	IGHV2-5	IGH4	AHRDVQVYVDS	11	IGLV2-8	IGJ2	SSYAGRDFL	11
2	PV18.A2.D3.P4.18	IgA2	IGHV3-23	IGH4	ESLVNSVYVMDV	12	IGHV4-1	IGK1	QDYSYPLT	8
3	PV18.G4.D3.P3.16	IgG4	IGHV4-4	IGH6	ARHLRFTGDMTPHHMDV	20	IGHV1-16	IGK2	QDYSYPLT	9
3	PV18.G4.D3.P3.4	IgG4	IGHV4-59	IGH6	ARHLRFTGDMTPHHMDV	20	IGHV1-16	IGK4	QDYSYPLT	9

**Table 3-4. Anti-Dsg Clones isolated by antibody phage display.** All clones were included if and only if they showed cell-surface staining by indirect immunofluorescence to confirm the PV phenotype. 96 clones in total were characterized, representing 80 different lineages. 6 lineages contain clones in more than one subclass. Clones from all libraries were produced and analyzed by a team of Christoph Ellebrecht, Eun Jung Choi, Shantanu Reddy, and Eric Mukherjee.

**A**

Lineages ( $\geq 5$ Reads)						
	Sample 1 Lineages	Sample 2 Lineages	Overlap	Total	Est Pop (Chapman)	% Coverage
PV17 IgA1	8478	8528	4687	12319	15426	79.86
PV17 IgA2	3804	3499	2096	5207	6351	81.99
PV17 IgG1	10965	13931	4610	20286	33133	61.23
PV17 IgG4	3436	4183	1727	5892	8322	70.80
PV18 IgA1	10917	11930	4504	18343	28915	63.44
PV18 IgA2	3974	4513	1626	6861	11028	62.21
PV18 IgG1	14834	6909	1824	19919	56170	35.46
PV18 IgG4	915	416	119	1212	3183	38.08

**B**

Lineages ( $\geq 5$ Reads)						
	Rep 1 Lineages	Rep 2 Lineages	Overlap	Total	Est Pop (Chapman)	% Coverage
PV16 IgA1	11818	9761	9540	12039	12093	99.56
PV16 IgA2	9892	7718	7418	10192	10293	99.02
PV16 IgG1	8526	12120	8326	12320	12412	99.26
PV16 IgG4	2346	2395	2291	2450	2453	99.86
PV17 IgA1 sample 1	7044	7233	6624	7653	7693	99.49
PV17 IgA2 sample 1	278	3793	278	3793	3794	99.97
PV17 IgG4 sample 1	2865	3145	2832	3178	3183	99.85
PV17 IgA1 sample 2	6374	7313	6135	7552	7599	99.38
PV17 IgA2 sample 2	3367	1258	1258	3367	3368	99.97
PV17 IgG1 sample 2	9737	9988	8502	11223	11440	98.10
PV17 IgG4 sample 2	4012	2924	2915	4021	4025	99.89
PV18 IgA1 sample 1	6778	7938	5957	8759	9033	96.97
PV18 IgA2 sample 1	3304	3368	3064	3608	3633	99.32
PV18 IgG1 sample 1	12297	12516	11273	13540	13654	99.17
PV18 IgA1 sample 2	8360	8733	7201	9892	10140	97.56
PV18 IgA2 sample 2	3759	3786	3448	4097	4128	99.24
PV18 IgG1 sample 2	6527	6555	6381	6701	6706	99.93

**Table 3-5. Capture-Recapture Estimates Across Biological and Technical Replicates.**

(A) For PV17 and PV18, two 30mL blood samples were drawn (biological replicates) and sequenced separately to determine depth of coverage and variability due to B-cell sampling. For samples with multiple technical replicates within a biological replicate, reads for a given biological replicate were pooled for the comparison.

(B) For several biological samples, RNA was sequenced twice (rep 1 and rep 2) to determine depth of sequencing. In all cases, capture-recapture estimates were performed for lineages containing equal to or greater than five reads.

Subclass	EC1	EC2	EC3	EC4	EC5
IgA1	6	2	2	1	2
IgA2	4	0	3	0	1
IgG1	8	0	1	3	1
IgG4	31	0	1	3	2

**Table 3-6. Summary of Epitope Mapping Experiments.** Scfv was used to pull down either Dsg1/2 or Dsg3/2 domain-swapped constructs to determine epitope specificity. Data is compiled across all four patients.

## **CHAPTER 4:**

### **Discussion and Future Directions**

#### **4.1 Summary of Findings**

We have elucidated the role of autoantibody subclass in pemphigus vulgaris by determining the effects of the constant region on variable region affinity, specificity, and blistering ability, as well as investigated the relationships between variable regions across different subclasses in order to determine the class-switch pathways are most relevant to disease. Our results have substantiated the idea that there are partially disjoint sub-populations within the set of anti-Dsg antibodies in PV, partitioned by preferred class-switch pathways – namely, a prominent IgA1 to IgA2 switch, a fairly common IgG1 to IgA1 switch, and a notable, but not as prominent, IgG4 to IgA2 switch. Notably minor was a direct switch connecting IgG1 and IgG4, the two subclasses thought to be most important for disease, and the two subclasses that have a reciprocal relationship during chronic antigen stimulation (which shifts from IgG1 to IgG4 in some cases). Because both IgG1 and IgG4 are present during active disease, and IgG1 predominates during remission, we had hypothesized that the IgG4 and IgG1 could be clonally related – our data instead suggests the idea that most anti-Dsg IgG1 and anti-Dsg IgG4 lineages do not overlap. This has implications not only on the pathogenesis of PV, but on the development of the IgG4 repertoire in general. These results can help understand IgG4-mediated autoimmune disease, allergen desensitization, and any other process, pathological or protective, resulting in IgG4 production.

While previous studies have shown that injection of anti-Dsg3 scFv or Fab was sufficient to cause blistering in model systems, there remained the possibility that the IgG1 or IgG4 constant region directly modulated the affinity or pathogenicity of an anti-Dsg antibody. By cloning several anti-Dsg3 variable regions onto both IgG1 and IgG4 constant regions, we managed to show that the constant region does not modulate the blister-inducing ability, Dsg3 internalization ability, or epitope preferences of pathogenic anti-Dsg3 antibodies. This suggests that a change in constant region from IgG1 to IgG4 alone is not sufficient to explain the pathogenic role of IgG4 autoantibodies in PV (see chapter 2).

This result buttressed further investigation of the lineage relationships between IgG1 and IgG4 in a panel of four PV patients (chapter 3). Through subclass-specific repertoire cloning, we have shown that the IgG4 anti-Dsg repertoire in PV mostly develops independently from the IgG1, sharing few lineages and few direct-switch relationships. This was an unexpected finding; given the reciprocal relationship between IgG1 and IgG4 over the course of disease in PV (IgG1 predominance during remission, IgG4 predominance during active disease) and the arrangement of constant region exons on the IgH locus, we expected to find that the majority of IgG4 clones switched from IgG1 during active disease. We have also characterized the relationship between IgA and IgG in PV for the first time, showing that IgA1, IgG1 and IgA2 often share clonal lineages with each other, and that IgA subclasses appear to switch from IgG subclasses (if they are related by direct switch at all), rather than the reverse. In particular, IgA1 and IgA2 in pemphigus are highly related, and some IgG1 share lineages with, or even



directly class switch to, IgA1. Also notable is the lack of detected lineage relationships between IgA1 to IgG4, casting doubt on the idea that the mucosal to mucocutaneous disease course in PV is due to cross-reactivity hiding in IgA1. More generally, the fact that switching seems to proceed in the IgG to IgA direction suggests that IgA is not important for the ontogeny of PV, but may instead be an epiphenomenon or serve a disease-modifying function, possibilities we discuss later in this chapter.

Our results suggest that there exist two major axes of class-switching in PV, tracing two different paths through the antibody repertoire. First there is an interconnected IgG1-IgA1-IgA2 repertoire reactive to a wide range of epitopes on the Dsg molecule, and a solitary IgG4 +/- IgA2 repertoire tightly focused on the N-terminal domains of Dsg. The implications of these newly-characterized lineage relationships in PV, and in other will be discussed in the next sections.

## **4.2 Discussion and Future Directions**

### **4.2.1 The Autoantibody Repertoire in PV is Highly Diverse**

For the first time, we have categorized subclass-specific repertoires in pemphigus vulgaris, managing to isolate an unprecedented 80 lineages worth of validated anti-Dsg3 and anti-Dsg1 antibodies. Particularly notable is the fact that the response is highly polyclonal, encompassing many different CDR3 sequences and VH genes; this stands in contrast with the fact that some antigen-specific sequences appear to be highly stereotyped and shared between patients (416). Repertoire studies of

dengue fever (275), influenza vaccination (417), and HIV (418) have shown convergent CDR sequences, and some autoantibodies show shared VH gene usage across patients (419)(420)(421). Previous antibody cloning efforts in PV identified shared usage of the VH1-46 gene segment between patients (367), and also found that some pathogenic anti-Dsg antibodies share a D/E-X-X-X-W motif (209). In contrast, we only found VH1-46 antibodies in two of our four patients, only one of which had VH1-46 IgG. Furthermore, there were no convergent CDR3 sequences between patients

Our results suggest that the response to desmoglein has the potential to be much more polyclonal than initially anticipated, across all subclasses studied. Despite this, the plurality of clones across the panel of patients seem to bind to the EC1 domain of Dsg; this EC1 preference is particularly notable in the IgG4 clones. The reasons for this are unclear; perhaps particular epitopes within the EC1 domain are more easily displayed on dendritic cells or other antigen-presenting cells, or more easily shed into the circulation and taken up by antigen-presenting cells after keratinocyte death (thus triggering an immune response), or are particularly proficient at triggering an IgG4 class switch pathway. More generally, it may still be the case that within each EC domain, different subclasses preferentially bind to different amino acid residues. Finer mapping using competition ELISA or hydrogen-deuterium exchange (422,423) may reveal that IgG4 clones segregate to different epitopes in the EC1 and EC2 domain than N-terminal-binding IgG1, IgA1, or IgA2 clones, possibly explaining the relationship between IgG4 and disease activity. It would be particularly interesting, for example, if IgG4 clones preferentially bound to residues critical for adhesion, while other subclasses did not.

## 4.2.2 The Potential Role of IgA in PV

The possibility of cross-reactivity to mucosal antigens in PV has previously been investigated by our laboratory in the context of the shared VH gene usage theory. Because VH1-46 antibodies with low somatic mutation counts were found in several PV patients (367), and the VH1-46 segment is also prominent in the response to rotavirus (424), it was postulated that VH1-46 antibodies could be cross-reactive to both Dsg3 and rotavirus coat protein; several such antibodies were isolated from a PV patient's IgM repertoire (401). The idea of an IgA origin of disease was also particularly intriguing because a substantial portion of the mucosal repertoire has been found to be polyreactive, suggesting that the tolerance mechanisms in the mucosal immune system are somewhat relaxed to help fend off infection and keep commensal bacteria in check (265,425,426). Our experiments have instead suggested that anti-Dsg IgG4 antibodies likely do not originate from IgA1. This is not entirely unexpected - the IgA1 to IgG4 pathway is relatively rare, representing 4% of all direct switches to IgG4 in the class-switching landscape outlined by the Quake group (137), and had not been characterized prior to that study. Furthermore, because IgA2 is distal to IgG4 on the IgH locus, IgA2 cannot be the origin of anti-Dsg IgG4.

Collectively, our experiments seem to demonstrate the opposite relationship – that of at least some IgA clones switching from IgG clones. There is a demonstrable parent-child relationship between some anti-Dsg IgG1 and some IgA1 in autoreactive B cell lineages in PV, suggesting that at least part of the anti-Dsg IgA1 repertoire arises directly from IgG1 through class switching. In vitro experiments have shown that CD40

engagement and TGFbeta can induce an IgM-IgG1-IgA sequential switch (410). In the Quake dataset, the IgG1 to IgA1 class-switch pathway is the second most common in the overall class-switching landscape, second only to the IgM to IgA1 switch (137).

Unexpectedly, our results also showed a close relationship between some anti-Dsg IgG4 and anti-Dsg IgA2 in PV. Such a relationship in the normal antibody repertoire was first discovered earlier this year in the Quake dataset; across a panel of 22 patients, totaling 120000 unique IgA2 sequences and 16000 unique IgG4 sequences, that study found a total of two sequences identical between IgG4 and IgA2, making our finding all the more remarkable (137). Because our protocols called for separately amplifying and sequencing each subclass using hinge-specific amplicons, our experiment is particularly well-powered to find potentially rare class-switch pathways like this one, compared to previous studies. Whether this IgG4-IgA2 axis is equally prevalent in normal individuals, or specifically upregulated in pathological states like PV or other IgG4-mediated conditions, is unknown, and could be the subject of future subclass-specific repertoire sequencing experiments.

It is also notable that the majority of characterized anti-Dsg IgA2-containing lineages show direct switching from IgA1, which often occurs via a well-characterized extrafollicular T-independent class-switch pathway induced by the secretion of APRIL from intestinal epithelial cells, triggered by antigens from mucosal normal flora (91). The possibility therefore exists that autoreactive IgA2 arises through both GC and extrafollicular responses. However, because we have not tracked the cell of origin

through its development, we cannot determine for sure which pathway gave rise to any given anti-Dsg clone.

Collectively, these experiments suggest that anti-Dsg IgA are not the origin of autoreactivity in PV, and that therefore the switch from IgG to IgA could simply be an epiphenomenon triggered by the localized effect of IgA class-switch mediators and/or cross-reactivity of anti-Dsg IgG clones to some unknown antigens present in the mucosa. However, it raises the possibility that an IgG to IgA switch could serve a disease-modifying function in PV. It would be interesting, for example, to determine by immunostaining whether the IgA deposits in the skin of patients with neutrophilic infiltrate belong to the IgA1, IgA2, or both subclasses. It would be even more interesting to determine whether the IgA-secreting B cells creating those deposits switched from IgG by repertoire cloning of IgA+ plasma cells. An IgG1 to IgA or IgG4 to IgA2 class-switch would then be shown to be contributing to the painful inflammation experienced by PV patients due to neutrophilic infiltrates.

#### **4.2.3 The Relationship Between Antigen-Specific IgG1 and IgG4 in PV**

The inverse relationship between IgG1 and IgG4 in the setting of chronic antigen stimulation is well-characterized, but subclass-specific repertoires have not extensively been analyzed in this context. The Quake dataset suggests that approximately 25% of all class switches leading to IgG4 originate in IgG1, while the plurality (38%) originate in IgG2 (137). Our results indicate that rather than a sequential switch, it appears to be the

case that anti-Dsg IgG1 and IgG4 exist in independent lineages, with the latter emerging during active disease.

A priori, there exist three possibilities for the relationship between anti-Dsg IgG1 and IgG4: direct switch from IgG1 to IgG4, independent switch from common precursors, or independent evolution. Of 39 anti-Dsg IgG4-containing lineages found by our study, 8 of them contain both IgG1 and IgG4 sequences, and only one of them displays an IgG1 to IgG4 direct switch. Taken at face value, this seems to suggest that the majority (31/39) of anti-Dsg IgG4 arises completely independently of IgG1, a small minority arises through direct switch (1/39), while the rest share common precursors with IgG1 without displaying a direct switch. In contrast, there exist 30 anti-Dsg IgA2-containing lineages; 21 of them contain IgA1 reads. Furthermore, of the 11 lineages containing IgA1 and IgA2 reads that we subjected to phylogenetic analysis, 10 of them showed a direct switch from IgA1 to IgA2.

Because we robustly detect overlap between two subclasses (IgA1 and IgA2) and by contrast, do not detect it in another pair (IgG1 and IgG4), our data suggests we are sufficiently powered to detect multi-subclass anti-Dsg lineages in PV. This suggests that at the level of our sampling, the IgG1 and IgG4 anti-Dsg clones are disjoint compared to IgA1 and IgA2; in particular, we have demonstrated that anti-Dsg IgG4 preferentially segregates into single-subclass over than multi-subclass lineages compared to the other three subclasses we have analyzed (see Figure 3-3D).

However, we cannot conclusively conclude that there do not exist IgG1 relatives for anti-Dsg IgG4 clones, for two overarching reasons. First, sufficient sampling of the B

cell repertoire by sequencing is an issue. While rarefaction analysis indicates that the repertoire sequencing has sampled over 70% of the lineages in the peripheral repertoire; however, we have only drawn 1% of the total peripheral blood, and only 2% of a human's B cells are circulating at any given time (46); there is always the possibility that the IgG1 relatives (either IgG1 precursors of IgG4 or IgG1 shared or of our detected anti-Dsg IgG4 clones simply have escaped sequencing, either because they are circulating and not captured or because they are residing in a secondary lymphoid organ and thus cannot be sampled by a blood draw. A repertoire-cloning experiment that sequences several more biological replicates of blood and obtains tissues from other organs (spleen, MALT, skin-draining lymph nodes) would improve the chances of capturing these putative IgG1 relatives. Related to this is the issue that, because the experiments sampled directly from whole PBMC RNA, we are probably over-representing transcripts from plasmablasts, which dilutes the contribution from memory B cells and therefore make them less likely to be captured by our amplification. To counteract this issue, future repertoire studies could bulk sort memory and plasmablasts separately (or sort single B cells), and prepare sequencing libraries from them independently, which would ensure that plasmablast RNA does not swamp out the contribution from memory B cells. Second, there is the possibility is that IgG1 precursors or relatives no longer exist; they may have existed only as transient intermediates during an IgM to IgG1 to IgG4 switch. Future experiments could solve this issue by longitudinally cloning subclass-specific repertoires during disease onset and

remission, though even this may not capture transient IgG1 intermediates if they don't circulate for sufficient duration.

There are also some drawbacks to the phage display technique worth considering as a source of error or undersampling. First is the issue of light chains, which presents several caveats. Even though we can screen between 100 million and 1 billion transformants per patient-subclass pair, the random pairing of heavy and light chains may have failed to pair some reactive heavy chains with compatible light chains. This leads to the possibility that there exists some related IgG1-IgG4 pair that wasn't sampled in either the IgG1 or IgG4 libraries, and therefore wasn't analyzed. This can be solved by constructing multiple libraries from a single patient-subclass pair and panning them independently, increasing the chances that a given anti-Dsg heavy chain pairs with a permissive light chain. Alternatively, a future study could use a technique that preserves heavy-light pairing (for example trapping individual cells in emulsified droplets (427)), perhaps coupled to a flow-cytometry-based technique for sorting single B cells for Dsg-reactivity; this would ensure that every heavy chain is screened while expressed alongside its cognate light chain, eliminating the question of whether an in vivo reactive heavy chain is missed.

There is also the possibility that some of our anti-Dsg antibodies primarily gain their reactivity from their light chain, rather than their heavy chains. While antibodies generally make most of their contacts through heavy chain CDR residues, there are examples in the literature where antigen recognition is encoded by the light chain (428)(429). In many cases, we found several anti-Dsg reactive heavy chains bound to



different light chains, suggesting that the heavy chain remains the most important determinant in binding in many cases; however, it may be the case that some of our anti-Dsg heavy chains actually have no anti-Dsg activity in vivo and are therefore irrelevant to disease, due to all anti-Dsg activity being mediated by the (mispaird) light chain. The methods outlined above, which preserve heavy-light pairing, would mitigate the possibility of irrelevant heavy chains entering the analysis.

Another overarching issue is that phage display is designed to find a particular type of antibody, rather than sample the full breadth of response to a given antigen. Phage display is optimized to find high-affinity clones, and in particular our stringent washing procedures may be selecting for high-affinity clones with slow disassociation rates. This has a few implications. First, high-affinity clones may be out-competing disease-relevant low-affinity clones binding to the same epitope, or simply due to high-affinity clones dominating growth during phage amplification. There may be entire lineages of disease-relevant low-affinity clones missed by our panning for this reason. Amplifying and panning the unbound phage from each round, or decreasing washing time, may pick up some of these low-affinity binders. Furthermore, it may be the case that disease-relevant antibodies derive their high affinity from short association rates, rather than long dissociation rates. Antibodies with short association rates could theoretically be found by decreasing the initial incubation time of the library, which would favor clones that quickly associated with Dsg.

Another drawback of our pipeline is that is that some positive clones may be toxic to phage or bacteria, meaning they fail to amplify during panning or may confer

growth disadvantages to the phage and will therefore be out-competed (430). Even small differences in growth rate in a library can lead to dropout of positive binders (431), suggesting that our screen may be missing several clones. This is further exacerbated by the fact that about half of the scFv that we attempted to produce in this experiment did not express in sufficient quantities for characterization, possibly due to aggregation or toxicity; this may be mitigated by expression and screening in eukaryotic systems. Another way to circumvent this issue is by using alternate screening technologies, like antigen-specific B cell sorting (432), yeast display (433), ribosome display (434), or mammalian cell surface display (435).

Given these caveats, it may be the case that we are missing certain relevant anti-Dsg antibodies and relatives in our samples. However, given the fact that our sampling, both through phage display and through sequencing, was sufficient to show that direct class-switching between anti-Dsg IgG1 and IgG4 in the periphery is not the predominant pathway of IgG4 production. The larger implications of this fact will be discussed in the next section.

#### **4.2.4 Mechanisms of PV Pathogenesis and Their Larger Implications**

The predominance of antigen-specific IgG4 in PV during active disease mirrors several other conditions, including other antibody-mediated autoimmune diseases (242) and subjects undergoing continuous antigen challenge, like allergen desensitization therapy (255,436), helminth infection (249)(437), and continuous exposure to bee venom antigens (248). In particular, an IgG1 to IgG4 shift occurs in beekeepers

undergoing repeated challenge by bee venom antigens; tolerant beekeepers produce IgG4 against bee venom from IL-10 secreting regulatory B cells, which dampen allergic reactions (407). Because IgG4 has no effector function, it is believed to serve as a “brake” on the immune response, competing with the binding of antibodies from other subclasses (253)(252). Further evidence for this fact includes that the regulatory cytokine IL-10 is responsible for biasing class-switch to IgG4 over IgE (438); this further suggests that the aforementioned IL-10 secreting B cells will bias other B cells to switch to IgG4 in the setting of chronic inflammation.

Because of the temporal relationships between IgG1 and IgG4 in endemic PV triggered by continuous exposure to sandfly antigens (404)(406)(439) and in beekeepers, as well as the presence of IgG1, but not IgG4, in some healthy relatives of PV patients (246), it is tempting to suggest that PV arises by a similar mechanism of chronic antigen stimulation, though rather than some external antigen it could be Dsg doing the stimulating. Perhaps, for example, continuous shedding of Dsg3 extracellular domain from apoptotic keratinocytes, which is then taken up and presented in skin-draining lymph nodes by Langerhans cells, cross-reactive B cells, or other antigen-presenting cells, eventually leads to a germinal center response in genetically susceptible individuals. Possibly suggestive of this, studies have found that polymorphisms in the pro-apoptotic molecule ST18 are associated with PV in some populations, and that ST18 is expressed at higher levels in the skin of PV patients compared to controls (440)(441). Also suggestive is the fact that HLA-susceptible individuals sometimes have circulating autoreactive T cells; in particular, some HLA-

susceptible individuals without disease seem to almost exclusively show autoreactive Th1 cells, while those with disease show both autoreactive Th1 and Th2 (442); this is concordant with the fact that IgG4 class-switch occurs in response to Th2 cytokines like IL-4 and IL-13 (443). This low-level stimulation over a lifetime would then trigger the switch to IgG4-secreting regulatory B cells, just like in beekeepers and the allergen-desensitized – however, rather than the IgG4 serving a protective role, it actually is the hallmark of active disease, causing blistering and pathogenesis rather than dampening the response as it has evolved to do. Repertoire cloning from regulatory B cell subsets could determine whether this portion of the mechanism is plausible; if, like in allergen desensitization therapy and in beekeepers, the anti-Dsg IgG4 is exclusively secreted by regulatory B cells, this would suggest a shared mechanism between these situations.

While the idea of continuous self-inoculation by Dsg resulting in disease is a plausible model, there are other possibilities for the shift to IgG4 that don't involve a direct response to chronic antigen stimulation. For example, a normally protective IgM B cell cross-reactive to both Dsg and some pathogen could be induced to switch to IgG4 in the presence of a permissive Th2 cytokine milieu, rather than responding to Dsg directly; this appears to be the mechanism in endemic PF, in which persistent challenge with sandfly antigens leads to IgG4 that is cross-reactive to Dsg1 (404). This possibility is particularly raised due to the cross-reactivity of VH1-46 IgM antibodies cloned from PV patients to both Dsg3 and the rotavirus protein VP6 (401); however, because we have determined that the anti-Dsg response is not VH restricted, such a single inciting event of cross-reactivity may not be sufficient to explain all instances of PV. It would be

intriguing, however, if cross-reactivity shaped and modified the response in some way. It is possible that, for example, IgG1 and IgG4 that arise from common precursors due to cross-reactivity of their IgM precursors to some pathogen (followed by a Th2 response that leads the same IgM B cell to switch independently to IgG1 and IgG4). Further testing of germline IgG4 and IgG1 precursors for polyreactivity or cross-reactivity to pathogens is necessary to determine whether this is a feasible outcome.

If PV, in fact, arises through the same mechanisms as other forms of chronic antigen stimulation, the largely disjoint lineage relationship between IgG1 and IgG4 we have characterized in PV suggests that chronic antigen stimulation could lead to the production of circulating IgG4 plasma cells directly from IgM, IgG2, or IgG3, which like IgG1 lie 5' to the IgG4 constant region on the IgH locus. These results also have implications for other IgG4-dominant autoimmune diseases and the IgG4 response to chronic antigen stimulation. It would be interesting, for example to longitudinally clone IgG1 and IgG4 in beekeepers to determine whether they share a lineage or arise independently, which would answer a fundamental question about the immune system's self-limiting response to constant antigen challenge. Determining the cell of origin, lineage, and maturation pathways of IgG4 in different types of chronic antigen stimulation, for example, could lead to therapeutics that specifically up-regulate (in the case where it's protective, i.e. allergen desensitization therapy) or down-regulate (in the case where it's pathogenic, i.e. PV and other IgG4-mediated autoimmune diseases) the IgG4 response in these situations, leading to more targeted therapies against

pathogenic IgG4-mediated conditions and superior responses to existing therapies that require an IgG4 response.

Confirming this idea of two separate switch pathways for the majority of IgG4 and the majority of IgG1 will require subclass-specific repertoire sequencing of IgG2 and IgG3 in PV patients; there remains a possibility that affinity maturation and class switching from these subclasses is the true origin of anti-Dsg IgG4, though this appears unlikely due to the lack of involvement of these subclasses in disease. Ruling those out at intermediates would then suggest that IgG4 originate directly from IgM. It would then be very interesting if in their germline configuration, anti-Dsg IgG4 clones show altered affinity, epitope coverage, and/or polyreactivity relative to anti-Dsg IgG1 germline sequences. Our experiments showed significant germline and early reactivity to Dsg in IgG4-containing lineages, suggesting that there is broader potential for naïve B cell anti-Dsg reactivity than previously appreciated, and that IgG4 may arise from such anti-Dsg IgM clones.

### **4.3 Summary**

This thesis has sought to answer the question of whether autoreactive IgG4 B cells arise from other autoreactive subclasses in PV. More broadly, we have asked the question: what are the clonal relationships between autoreactive B cells belonging to different subclasses in PV? We have provided compelling evidence that autoreactive IgG4 clones in PV largely do not share lineages with other subclasses, suggesting that

IgG4 in other settings may also arise independently. This has implications on the pathogenesis of other IgG4-mediated autoimmune disease and on the mechanism behind the response to chronic antigen stimulation in general. Future experiments should focus on generalizing this result to include more patients and more subclasses, on testing the implications of this result for IgG4 responses in the setting of chronic antigen stimulation or autoimmune disease, and on further characterization of germline and somatically mutated intermediates in established anti-Dsg lineages to better understand the evolution of the autoreactive B cell repertoire in PV and the IgG4 repertoire in general.

## **CHAPTER 5:**

### **References**

1. Borges JL, Giral A. *The Library of Babel*. Boston: David R. Godine Publisher; 2000. 39 p.
2. Dennett DC. *Darwin's Dangerous Idea: Evolution and the Meaning of Life*. Reprint edition. Simon & Schuster; 2014. 594 p.
3. Ostermeier M. Beyond cataloging the Library of Babel. *Chem Biol*. 2007 Mar;14(3):237–8.
4. Arnold FH. “The Library of Maynard-Smith: My Search for Meaning in the protein universe”. *Adv Protein Chem*. 2000;55:ix–xi.
5. Autoimmune Statistics [Internet]. AARDA. [cited 2016 Sep 7]. Available from: <https://www.aarda.org/autoimmune-information/autoimmune-statistics/>
6. Hirano M, Das S, Guo P, Cooper MD. The evolution of adaptive immunity in vertebrates. *Adv Immunol*. 2011;109:125–57.
7. Matsuda F, Ishii K, Bourvagnet P, Kuma K, Hayashida H, Miyata T, et al. The Complete Nucleotide Sequence of the Human Immunoglobulin Heavy Chain Variable Region Locus. *J Exp Med*. 1998 Dec 7;188(11):2151–62.
8. McBride OW, Hieter PA, Hollis GF, Swan D, Otey MC, Leder P. Chromosomal location of human kappa and lambda immunoglobulin light chain constant region genes. *J Exp Med*. 1982 May 1;155(5):1480–90.
9. Krangel MS. The Ties that Bind (the Igh Locus). *Trends Genet TIG*. 2016 May;32(5):253–5.
10. Tonegawa S. Somatic generation of antibody diversity. *Nature*. 1983 Apr 14;302(5909):575–81.
11. Hozumi N, Tonegawa S. Evidence for somatic rearrangement of immunoglobulin genes coding for variable and constant regions. *Proc Natl Acad Sci U S A*. 1976 Oct;73(10):3628–32.
12. Nutt SL, Heavey B, Rolink AG, Busslinger M. Commitment to the B-lymphoid lineage depends on the transcription factor Pax5. *Nature*. 1999 Dec 16;402:14–20.
13. Lee B, Dekker JD, Lee B, Iyer VR, Sleckman BP, Shaffer AL, et al. The BCL11A Transcription Factor Directly Activates RAG Gene Expression and V(D)J Recombination. *Mol Cell Biol*. 2013 May 1;33(9):1768–81.



14. Oettinger MA, Schatz DG, Gorka C, Baltimore D. RAG-1 and RAG-2, adjacent genes that synergistically activate V(D)J recombination. *Science*. 1990 Jun 22;248(4962):1517–23.
15. McBlane JF, van Gent DC, Ramsden DA, Romeo C, Cuomo CA, Gellert M, et al. Cleavage at a V(D)J recombination signal requires only RAG1 and RAG2 proteins and occurs in two steps. *Cell*. 1995 Nov 3;83(3):387–95.
16. Kim M-S, Lapkouski M, Yang W, Gellert M. Crystal Structure of the V(D)J Recombinase RAG1-RAG2. *Nature*. 2015 Feb 26;518(7540):507–11.
17. Ru H, Chambers MG, Fu T-M, Tong AB, Liao M, Wu H. Molecular Mechanism of V(D)J Recombination from Synaptic RAG1-RAG2 Complex Structures. *Cell*. 2015 Nov 19;163(5):1138–52.
18. Lapkouski M, Chuenchor W, Kim M-S, Gellert M, Yang W. Assembly Pathway and Characterization of the RAG1/2-DNA Paired and Signal-end Complexes. *J Biol Chem*. 2015 Jun 5;290(23):14618–25.
19. Yaoita Y, Matsunami N, Choi CY, Sugiyama H, Kishimoto T, Honjo T. The D-JH complex is an intermediate to the complete immunoglobulin heavy-chain V-region gene. *Nucleic Acids Res*. 1983 Nov 11;11(21):7303–16.
20. Baltimore D. Is terminal deoxynucleotidyl transferase a somatic mutagen in lymphocytes? *Nature*. 1974 Mar 29;248(447):409–11.
21. Desiderio SV, Yancopoulos GD, Paskind M, Thomas E, Boss MA, Landau N, et al. Insertion of N regions into heavy-chain genes is correlated with expression of terminal deoxytransferase in B cells. *Nature*. 1984 Oct 25;311(5988):752–5.
22. Motea EA, Berdis AJ. Terminal Deoxynucleotidyl Transferase: The Story of a Misguided DNA Polymerase. *Biochim Biophys Acta*. 2010 May;1804(5):1151–66.
23. Chang HHY, Lieber MR. Structure-Specific nuclease activities of Artemis and the Artemis: DNA-PKcs complex. *Nucleic Acids Res*. 2016 Jun 20;44(11):4991–7.
24. Malu S, De Ioannes P, Kozlov M, Greene M, Francis D, Hanna M, et al. Artemis C-terminal region facilitates V(D)J recombination through its interactions with DNA Ligase IV and DNA-PKcs. *J Exp Med*. 2012 May 7;209(5):955–63.
25. Jackson KJ, Gaeta B, Sewell W, Collins AM. Exonuclease activity and P nucleotide addition in the generation of the expressed immunoglobulin repertoire. *BMC Immunol*. 2004;5:19.
26. Grawunder U, Zimmer D, Fugmann S, Schwarz K, Lieber MR. DNA Ligase IV Is Essential for V(D)J Recombination and DNA Double-Strand Break Repair in Human Precursor Lymphocytes. *Mol Cell*. 1998 Oct;2(4):477–84.

27. Malu S, Malshetty V, Francis D, Cortes P. Role of non-homologous end joining in V(D)J recombination. *Immunol Res.* 2012 Dec;54(1–3):233–46.
28. Alt FW, Yancopoulos GD, Blackwell TK, Wood C, Thomas E, Boss M, et al. Ordered rearrangement of immunoglobulin heavy chain variable region segments. *EMBO J.* 1984 Jun;3(6):1209–19.
29. Zhang M, Srivastava G, Lu L. The pre-B cell receptor and its function during B cell development. *Cell Mol Immunol.* 2004 Apr;1(2):89–94.
30. Mårtensson I-L, Almqvist N, Grimsholm O, Bernardi AI. The pre-B cell receptor checkpoint. *FEBS Lett.* 2010 Jun 18;584(12):2572–9.
31. Anbazhagan K, Rabbind Singh A, Isabelle P, Stella I, Céline A-DM, Bissac E, et al. Human pre-B cell receptor signal transduction: evidence for distinct roles of PI3kinase and MAP-kinase signalling pathways. *Immun Inflamm Dis.* 2013 Oct;1(1):26–36.
32. Melchers F, ten Boekel E, Seidl T, Kong XC, Yamagami T, Onishi K, et al. Repertoire selection by pre-B-cell receptors and B-cell receptors, and genetic control of B-cell development from immature to mature B cells. *Immunol Rev.* 2000 Jun;175:33–46.
33. Bergman Y, Cedar H. A stepwise epigenetic process controls immunoglobulin allelic exclusion. *Nat Rev Immunol.* 2004 Oct;4(10):753–61.
34. Mårtensson I-L, Rolink A, Melchers F, Mundt C, Licence S, Shimizu T. The pre-B cell receptor and its role in proliferation and Ig heavy chain allelic exclusion. *Semin Immunol.* 2002 Oct;14(5):335–42.
35. Vettermann C, Schlissel MS. Allelic exclusion of immunoglobulin genes: models and mechanisms. *Immunol Rev.* 2010 Sep;237(1):22–42.
36. Nishimoto N, Kubagawa H, Ohno T, Gartland GL, Stankovic AK, Cooper MD. Normal pre-B cells express a receptor complex of mu heavy chains and surrogate light-chain proteins. *Proc Natl Acad Sci U S A.* 1991 Jul 15;88(14):6284–8.
37. Melchers F. Checkpoints that control B cell development. *J Clin Invest.* 2015 Jun;125(6):2203–10.
38. Boyd SD, Gaëta BA, Jackson KJ, Fire AZ, Marshall EL, Merker JD, et al. Individual variation in the germline Ig gene repertoire inferred from variable region gene rearrangements. *J Immunol Baltim Md 1950.* 2010 Jun 15;184(12):6986–92.
39. Lefranc MP. Nomenclature of the human immunoglobulin heavy (IGH) genes. *Exp Clin Immunogenet.* 2001;18(2):100–16.

40. Lefranc MP. IMGT, the international ImMunoGeneTics database. *Nucleic Acids Res.* 2001 Jan 1;29(1):207–9.
41. Jr CAJ, Travers P, Walport M, Shlomchik MJ, Jr CAJ, Travers P, et al. *Immunobiology*. 5th ed. Garland Science; 2001.
42. Lefranc MP. Nomenclature of the human immunoglobulin lambda (IGL) genes. *Exp Clin Immunogenet.* 2001;18(4):242–54.
43. Lefranc MP. Nomenclature of the human immunoglobulin kappa (IGK) genes. *Exp Clin Immunogenet.* 2001;18(3):161–74.
44. Williams SC, Fripiat JP, Tomlinson IM, Ignatovich O, Lefranc MP, Winter G. Sequence and evolution of the human germline V lambda repertoire. *J Mol Biol.* 1996 Nov 29;264(2):220–32.
45. Dimitrov DS. Therapeutic antibodies, vaccines and antibodyomes. *mAbs.* 2010;2(3):347–56.
46. Georgiou G, Ippolito GC, Beausang J, Busse CE, Wardemann H, Quake SR. The promise and challenge of high-throughput sequencing of the antibody repertoire. *Nat Biotechnol.* 2014 Feb;32(2):158–68.
47. Yaari G, Kleinstein SH. Practical guidelines for B-cell receptor repertoire sequencing analysis. *Genome Med* [Internet]. 2015 Nov 20 [cited 2016 Sep 8];7. Available from: <http://www.ncbi.nlm.nih.gov/pmc/articles/PMC4654805/>
48. Rajewsky K. Clonal selection and learning in the antibody system. *Nature.* 1996 Jun 27;381(6585):751–8.
49. Ehlich A, Martin V, Müller W, Rajewsky K. Analysis of the B-cell progenitor compartment at the level of single cells. *Curr Biol.* 1994 Jul;4(7):573–83.
50. Prak EL, Weigert M. Light chain replacement: a new model for antibody gene rearrangement. *J Exp Med.* 1995 Aug 1;182(2):541–8.
51. Sela-Culang I, Kunik V, Ofan Y. The structural basis of antibody-antigen recognition. *B Cell Biol.* 2013;4:302.
52. Al-Lazikani B, Lesk AM, Chothia C. Standard conformations for the canonical structures of immunoglobulins. *J Mol Biol.* 1997 Nov 7;273(4):927–48.
53. Jones PT, Dear PH, Foote J, Neuberger MS, Winter G. Replacing the complementarity-determining regions in a human antibody with those from a mouse. *Nature.* 1986 Jun 29;321(6069):522–5.
54. MacCallum RM, Martin ACR, Thornton JM. Antibody-antigen Interactions: Contact Analysis and Binding Site Topography. *J Mol Biol.* 1996 Oct 11;262(5):732–45.

55. Wu TT, Kabat EA. An analysis of the sequences of the variable regions of Bence Jones proteins and myeloma light chains and their implications for antibody complementarity. *J Exp Med*. 1970 Aug 1;132(2):211–50.
56. Padlan EA. Anatomy of the antibody molecule. *Mol Immunol*. 1994 Feb;31(3):169–217.
57. Robin G, Sato Y, Desplancq D, Rochel N, Weiss E, Martineau P. Restricted Diversity of Antigen Binding Residues of Antibodies Revealed by Computational Alanine Scanning of 227 Antibody–Antigen Complexes. *J Mol Biol*. 2014 Nov 11;426(22):3729–43.
58. Klein F, Diskin R, Scheid JF, Gaebler C, Mouquet H, Georgiev IS, et al. Somatic mutations of the immunoglobulin framework are generally required for broad and potent HIV-1 neutralization. *Cell*. 2013 Mar 28;153(1):126–38.
59. Hershberg U, Luning Prak ET. The analysis of clonal expansions in normal and autoimmune B cell repertoires. *Philos Trans R Soc B Biol Sci [Internet]*. 2015 Sep 5 [cited 2016 Sep 7];370(1676). Available from: <http://www.ncbi.nlm.nih.gov/pmc/articles/PMC4528416/>
60. Chung JB, Silverman M, Monroe JG. Transitional B cells: step by step towards immune competence. *Trends Immunol*. 2003 Jun 1;24(6):342–8.
61. Agrawal S, Smith SABC, Tangye SG, Sewell WA. Transitional B cell subsets in human bone marrow. *Clin Exp Immunol*. 2013 Oct;174(1):53–9.
62. Palanichamy A, Barnard J, Zheng B, Owen T, Quach T, Wei C, et al. Novel human transitional B cell populations revealed by B cell depletion therapy. *J Immunol Baltim Md 1950*. 2009 May 15;182(10):5982–93.
63. De Silva NS, Klein U. Dynamics of B cells in germinal centres. *Nat Rev Immunol*. 2015 Mar;15(3):137–48.
64. Avalos AM, Ploegh HL. Early BCR Events and Antigen Capture, Processing, and Loading on MHC Class II on B Cells. *Front Immunol [Internet]*. 2014 Mar 10 [cited 2016 Sep 13];5. Available from: <http://www.ncbi.nlm.nih.gov/pmc/articles/PMC3948085/>
65. Batista FD, Harwood NE. The who, how and where of antigen presentation to B cells. *Nat Rev Immunol*. 2009 Jan;9(1):15–27.
66. Laydon DJ, Bangham CRM, Asquith B. Estimating T-cell repertoire diversity: limitations of classical estimators and a new approach. *Philos Trans R Soc B Biol Sci [Internet]*. 2015 Aug 19 [cited 2016 Sep 19];370(1675). Available from: <http://www.ncbi.nlm.nih.gov/pmc/articles/PMC4528489/>

67. Qi H, Cannons JL, Klauschen F, Schwartzberg PL, Germain RN. SAP-controlled T-B cell interactions underlie germinal centre formation. *Nature*. 2008 Oct 9;455(7214):764–9.
68. Okada T, Miller MJ, Parker I, Krummel MF, Neighbors M, Hartley SB, et al. Antigen-engaged B cells undergo chemotaxis toward the T zone and form motile conjugates with helper T cells. *PLoS Biol*. 2005 Jun;3(6):e150.
69. Lederman S, Yellin MJ, Inghirami G, Lee JJ, Knowles DM, Chess L. Molecular interactions mediating T-B lymphocyte collaboration in human lymphoid follicles. Roles of T cell-B-cell-activating molecule (5c8 antigen) and CD40 in contact-dependent help. *J Immunol Baltim Md 1950*. 1992 Dec 15;149(12):3817–26.
70. Slavik JM, Hutchcroft JE, Bierer BE. CD28/CTLA-4 and CD80/CD86 families: signaling and function. *Immunol Res*. 1999;19(1):1–24.
71. Pentcheva-Hoang T, Egen JG, Wojnoonski K, Allison JP. B7-1 and B7-2 selectively recruit CTLA-4 and CD28 to the immunological synapse. *Immunity*. 2004 Sep;21(3):401–13.
72. Bromley SK, Iaboni A, Davis SJ, Whitty A, Green JM, Shaw AS, et al. The immunological synapse and CD28-CD80 interactions. *Nat Immunol*. 2001 Dec;2(12):1159–66.
73. Chandra V, Bortnick A, Murre C. AID targeting: old mysteries and new challenges. *Trends Immunol*. 2015 Sep;36(9):527–35.
74. Ramiro AR, Stavropoulos P, Jankovic M, Nussenzweig MC. Transcription enhances AID-mediated cytidine deamination by exposing single-stranded DNA on the nontemplate strand. *Nat Immunol*. 2003 May;4(5):452–6.
75. Chaudhuri J, Tian M, Khuong C, Chua K, Pinaud E, Alt FW. Transcription-targeted DNA deamination by the AID antibody diversification enzyme. *Nature*. 2003 Apr 17;422(6933):726–30.
76. Dickerson SK, Market E, Besmer E, Papavasiliou FN. AID mediates hypermutation by deaminating single stranded DNA. *J Exp Med*. 2003 May 19;197(10):1291–6.
77. Schanz S, Castor D, Fischer F, Jiricny J. Interference of mismatch and base excision repair during the processing of adjacent U/G mispairs may play a key role in somatic hypermutation. *Proc Natl Acad Sci U S A*. 2009 Apr 7;106(14):5593–8.
78. Saribasak H, Gearhart PJ. Does DNA repair occur during somatic hypermutation? *Semin Immunol*. 2012 Aug;24(4):287–92.
79. Roa S, Li Z, Peled JU, Zhao C, Edelmann W, Scharff MD. MSH2/MSH6 complex promotes error-free repair of AID-induced dU:G mispairs as well as error-prone hypermutation of A:T sites. *PloS One*. 2010;5(6):e11182.

80. Martin A, Scharff MD. AID and mismatch repair in antibody diversification. *Nat Rev Immunol.* 2002 Aug;2(8):605–14.
81. Wilson TM, Vaisman A, Martomo SA, Sullivan P, Lan L, Hanaoka F, et al. MSH2-MSH6 stimulates DNA polymerase  $\eta$ , suggesting a role for A:T mutations in antibody genes. *J Exp Med.* 2005 Feb 21;201(4):637–45.
82. Zeng X, Winter DB, Kasmer C, Kraemer KH, Lehmann AR, Gearhart PJ. DNA polymerase  $\eta$  is an A-T mutator in somatic hypermutation of immunoglobulin variable genes. *Nat Immunol.* 2001 Jun;2(6):537–41.
83. Meyer-Hermann ME, Maini PK, Iber D. An analysis of B cell selection mechanisms in germinal centers. *Math Med Biol J IMA.* 2006 Sep;23(3):255–77.
84. Zhang Y, Garcia-Ibanez L, Toellner K-M. Regulation of germinal center B-cell differentiation. *Immunol Rev.* 2016 Mar 1;270(1):8–19.
85. Tangye SG. Staying alive: regulation of plasma cell survival. *Trends Immunol.* 2011 Dec 1;32(12):595–602.
86. Nutt SL, Hodgkin PD, Tarlinton DM, Corcoran LM. The generation of antibody-secreting plasma cells. *Nat Rev Immunol.* 2015 Mar;15(3):160–71.
87. Kurosaki T, Kometani K, Ise W. Memory B cells. *Nat Rev Immunol.* 2015 Mar;15(3):149–59.
88. Kruetzmann S, Rosado MM, Weber H, Germing U, Tournilhac O, Peter H-H, et al. Human Immunoglobulin M Memory B Cells Controlling *Streptococcus pneumoniae* Infections Are Generated in the Spleen. *J Exp Med.* 2003 Apr 7;197(7):939–45.
89. Rawlings DJ, Schwartz MA, Jackson SW, Meyer-Bahlburg A. Integration of B cell responses through Toll-like receptors and antigen receptors. *Nat Rev Immunol.* 2012 Apr;12(4):282–94.
90. Litinskiy MB, Nardelli B, Hilbert DM, He B, Schaffer A, Casali P, et al. DCs induce CD40-independent immunoglobulin class switching through BLyS and APRIL. *Nat Immunol.* 2002 Sep;3(9):822–9.
91. He B, Xu W, Santini PA, Polydorides AD, Chiu A, Estrella J, et al. Intestinal bacteria trigger T cell-independent immunoglobulin A(2) class switching by inducing epithelial-cell secretion of the cytokine APRIL. *Immunity.* 2007 Jun;26(6):812–26.
92. Paus D, Phan TG, Chan TD, Gardam S, Basten A, Brink R. Antigen recognition strength regulates the choice between extrafollicular plasma cell and germinal center B cell differentiation. *J Exp Med.* 2006 Apr 17;203(4):1081–91.
93. McHeyzer-Williams MG, McLean MJ, Lalor PA, Nossal GJ. Antigen-driven B cell differentiation in vivo. *J Exp Med.* 1993 Jul 1;178(1):295–307.

94. Belessi C, Stamatopoulos K, Stavroyianni N, Zoi K, Papadaki T, Kosmas C. Somatic hypermutation targeting to intrinsic hotspots of immunoglobulin genes in follicular lymphoma and multiple myeloma. *Leukemia*. 2001 Nov;15(11):1772–8.
95. Milstein C, Neuberger MS, Staden R. Both DNA strands of antibody genes are hypermutation targets. *Proc Natl Acad Sci U S A*. 1998 Jul 21;95(15):8791–4.
96. Rogozin IB, Kolchanov NA. Somatic hypermutagenesis in immunoglobulin genes. II. Influence of neighbouring base sequences on mutagenesis. *Biochim Biophys Acta*. 1992 Nov 15;1171(1):11–8.
97. Rogozin IB, Diaz M. Cutting Edge: DGYW/WRCH Is a Better Predictor of Mutability at G:C Bases in Ig Hypermutation Than the Widely Accepted RGYW/WRCY Motif and Probably Reflects a Two-Step Activation-Induced Cytidine Deaminase-Triggered Process. *J Immunol*. 2004 Mar 15;172(6):3382–4.
98. Rogozin IB, Pavlov YI, Bebenek K, Matsuda T, Kunkel TA. Somatic mutation hotspots correlate with DNA polymerase  $\eta$  error spectrum. *Nat Immunol*. 2001 Jun;2(6):530–6.
99. Longo NS, Satorius CL, Plebani A, Durandy A, Lipsky PE. Characterization of immunoglobulin gene somatic hypermutation in the absence of activation-induced cytidine deaminase. *J Immunol Baltim Md 1950*. 2008 Jul 15;181(2):1299–306.
100. Yavuz S, Yavuz AS, Kraemer KH, Lipsky PE. The role of polymerase eta in somatic hypermutation determined by analysis of mutations in a patient with xeroderma pigmentosum variant. *J Immunol Baltim Md 1950*. 2002 Oct 1;169(7):3825–30.
101. Yaari G, Vander Heiden JA, Uduman M, Gadala-Maria D, Gupta N, Stern JNH, et al. Models of Somatic Hypermutation Targeting and Substitution Based on Synonymous Mutations from High-Throughput Immunoglobulin Sequencing Data. *Front Immunol [Internet]*. 2013 Nov 15 [cited 2016 Sep 7];4. Available from: <http://www.ncbi.nlm.nih.gov/pmc/articles/PMC3828525/>
102. Jolly CJ, Wagner SD, Rada C, Klix N, Milstein C, Neuberger MS. The targeting of somatic hypermutation. *Semin Immunol*. 1996 Jun 1;8(3):159–68.
103. Wei L, Chahwan R, Wang S, Wang X, Pham PT, Goodman MF, et al. Overlapping hotspots in CDRs are critical sites for V region diversification. *Proc Natl Acad Sci*. 2015 Feb 17;112(7):E728–37.
104. Wagner SD, Milstein C, Neuberger MS. Codon bias targets mutation. *Nature*. 1995 Aug 31;376(6543):732.
105. Saini J, Hershberg U. B cell variable genes have evolved their codon usage to focus the targeted patterns of somatic mutation on the complementarity determining regions. *Mol Immunol*. 2015 May;65(1):157–67.

106. Wilson PC, Bouteiller O de, Liu Y-J, Potter K, Banchereau J, Capra JD, et al. Somatic Hypermutation Introduces Insertions and Deletions into Immunoglobulin V Genes. *J Exp Med*. 1998 Jan 5;187(1):59–70.
107. Bowers PM, Verdino P, Wang Z, da Silva Correia J, Chhoa M, Macondray G, et al. Nucleotide Insertions and Deletions Complement Point Mutations to Massively Expand the Diversity Created by Somatic Hypermutation of Antibodies. *J Biol Chem*. 2014 Nov 28;289(48):33557–67.
108. Reason DC, Zhou J. Codon insertion and deletion functions as a somatic diversification mechanism in human antibody repertoires. *Biol Direct*. 2006;1:24.
109. Briney BS, Willis JR, Crowe JE. Location and length distribution of somatic hypermutation-associated DNA insertions and deletions reveals regions of antibody structural plasticity. *Genes Immun*. 2012 Oct;13(7):523–9.
110. de Wildt RMT, van Venrooij WJ, Winter G, Hoet RMA, Tomlinson IM. Somatic insertions and deletions shape the human antibody repertoire. *J Mol Biol*. 1999 Dec 3;294(3):701–10.
111. Krause JC, Tumpey TM, Huffman CJ, McGraw PA, Pearce MB, Tsibane T, et al. Naturally occurring human monoclonal antibodies neutralize both 1918 and 2009 pandemic influenza A (H1N1) viruses. *J Virol*. 2010 Mar;84(6):3127–30.
112. Yu X, Tsibane T, McGraw PA, House FS, Keefer CJ, Hicar MD, et al. Neutralizing antibodies derived from the B cells of 1918 influenza pandemic survivors. *Nature*. 2008 Sep 25;455(7212):532–6.
113. Xu R, Ekiert DC, Krause JC, Hai R, Crowe JE, Wilson IA. Structural basis of preexisting immunity to the 2009 H1N1 pandemic influenza virus. *Science*. 2010 Apr 16;328(5976):357–60.
114. Kepler TB, Liao H-X, Alam SM, Bhaskarabhatla R, Zhang R, Yandava C, et al. Immunoglobulin gene insertions and deletions in the affinity maturation of HIV-1 broadly reactive neutralizing antibodies. *Cell Host Microbe*. 2014 Sep 10;16(3):304–13.
115. Tan J, Pieper K, Piccoli L, Abdi A, Foglierini M, Geiger R, et al. A LAIR1 insertion generates broadly reactive antibodies against malaria variant antigens. *Nature*. 2016 Jan 7;529(7584):105–9.
116. Gupta NT, Vander Heiden JA, Uduman M, Gadala-Maria D, Yaari G, Kleinstein SH. Change-O: a toolkit for analyzing large-scale B cell immunoglobulin repertoire sequencing data. *Bioinforma Oxf Engl*. 2015 Oct 15;31(20):3356–8.
117. Chaudhuri J, Basu U, Zarrin A, Yan C, Franco S, Perlot T, et al. Evolution of the immunoglobulin heavy chain class switch recombination mechanism. *Adv Immunol*. 2007;94:157–214.



118. Flanagan JG, Rabbitts TH. Arrangement of human immunoglobulin heavy chain constant region genes implies evolutionary duplication of a segment containing gamma, epsilon and alpha genes. *Nature*. 1982 Dec 23;300(5894):709–13.
119. Stavnezer J, Schrader CE. IgH chain class switch recombination: mechanism and regulation. *J Immunol Baltim Md 1950*. 2014 Dec 1;193(11):5370–8.
120. Kinoshita K, Harigai M, Fagarasan S, Muramatsu M, Honjo T. A hallmark of active class switch recombination: Transcripts directed by I promoters on looped-out circular DNAs. *Proc Natl Acad Sci U S A*. 2001 Oct 23;98(22):12620–3.
121. Geisberger R, Lamers M, Achatz G. The riddle of the dual expression of IgM and IgD. *Immunology*. 2006 Aug;118(4):429–37.
122. Pioli PD, Debnath I, Weis JJ, Weis JH. Zfp318 regulates IgD expression by abrogating transcription termination within the Ighm/Ighd locus. *J Immunol Baltim Md 1950*. 2014 Sep 1;193(5):2546–53.
123. Enders A, Short A, Miosge LA, Bergmann H, Sontani Y, Bertram EM, et al. Zinc-finger protein ZFP318 is essential for expression of IgD, the alternatively spliced Igh product made by mature B lymphocytes. *Proc Natl Acad Sci U S A*. 2014 Mar 25;111(12):4513–8.
124. Otero DC, Rickert RC. CD19 Function in Early and Late B Cell Development. II. CD19 Facilitates the Pro-B/Pre-B Transition. *J Immunol*. 2003 Dec 1;171(11):5921–30.
125. Cariappa A, Chase C, Liu H, Russell P, Pillai S. Naive recirculating B cells mature simultaneously in the spleen and bone marrow. *Blood*. 2007 Mar 15;109(6):2339–45.
126. Revy P, Muto T, Levy Y, Geissmann F, Plebani A, Sanal O, et al. Activation-Induced Cytidine Deaminase (AID) Deficiency Causes the Autosomal Recessive Form of the Hyper-IgM Syndrome (HIGM2). *Cell*. 2000 Sep 1;102(5):565–75.
127. Muramatsu M, Kinoshita K, Fagarasan S, Yamada S, Shinkai Y, Honjo T. Class Switch Recombination and Hypermutation Require Activation-Induced Cytidine Deaminase (AID), a Potential RNA Editing Enzyme. *Cell*. 2000 Sep 1;102(5):553–63.
128. Longacre A, Storb U. A Novel Cytidine Deaminase Affects Antibody Diversity. *Cell*. 2000 Sep 1;102(5):541–4.
129. Shimizu A, Honjo T. Immunoglobulin class switching. *Cell*. 1984 Apr 1;36(4):801–3.
130. Cameron L, Gounni AS, Frenkiel S, Lavigne F, Vercelli D, Hamid Q. S $\epsilon$ S $\mu$  and S $\epsilon$ S $\gamma$  Switch Circles in Human Nasal Mucosa Following Ex Vivo Allergen

- Challenge: Evidence for Direct as Well as Sequential Class Switch Recombination. *J Immunol*. 2003 Oct 1;171(7):3816–22.
131. Xiong H, Dolpady J, Wabl M, Curotto de Lafaille MA, Lafaille JJ. Sequential class switching is required for the generation of high affinity IgE antibodies. *J Exp Med*. 2012 Feb 13;209(2):353–64.
  132. Mandler R, Finkelman FD, Levine AD, Snapper CM. IL-4 induction of IgE class switching by lipopolysaccharide-activated murine B cells occurs predominantly through sequential switching. *J Immunol Baltim Md 1950*. 1993 Jan 15;150(2):407–18.
  133. Looney TJ, Lee J-Y, Roskin KM, Hoh RA, King J, Glanville J, et al. Human B-cell isotype switching origins of IgE. *J Allergy Clin Immunol*. 2016 Feb;137(2):579–586.e7.
  134. van Zelm MC. B cells take their time: sequential IgG class switching over the course of an immune response? *Immunol Cell Biol*. 2014 Sep;92(8):645–6.
  135. Jackson KJL, Wang Y, Collins AM. Human immunoglobulin classes and subclasses show variability in VDJ gene mutation levels. *Immunol Cell Biol*. 2014 Sep;92(8):729–33.
  136. Berkowska MA, Driessen GJA, Bikos V, Grosserichter-Wagener C, Stamatopoulos K, Cerutti A, et al. Human memory B cells originate from three distinct germinal center-dependent and -independent maturation pathways. *Blood*. 2011 Aug 25;118(8):2150–8.
  137. Horns F, Vollmers C, Croote D, Mackey SF, Swan GE, Dekker CL, et al. Lineage tracing of human B cells reveals the in vivo landscape of human antibody class switching. *eLife*. 2016 Aug 2;5:e16578.
  138. Wardemann H, Nussenzweig MC. B-cell self-tolerance in humans. *Adv Immunol*. 2007;95:83–110.
  139. Shlomchik MJ. Sites and stages of autoreactive B cell activation and regulation. *Immunity*. 2008 Jan;28(1):18–28.
  140. Wardemann H, Yurasov S, Schaefer A, Young JW, Meffre E, Nussenzweig MC. Predominant Autoantibody Production by Early Human B Cell Precursors. *Science*. 2003 Sep 5;301(5638):1374–7.
  141. Kinnunen T, Chamberlain N, Morbach H, Cantaert T, Lynch M, Preston-Hurlburt P, et al. Specific peripheral B cell tolerance defects in patients with multiple sclerosis. *J Clin Invest*. 2013 Jun;123(6):2737–41.

142. Samuels J, Ng Y-S, Coupillaud C, Paget D, Meffre E. Impaired early B cell tolerance in patients with rheumatoid arthritis. *J Exp Med*. 2005 May 16;201(10):1659–67.
143. Lee J, Stathopoulos P, Gupta S, Bannock JM, Barohn RJ, Cotzomi E, et al. Compromised fidelity of B-cell tolerance checkpoints in AChR and MuSK myasthenia gravis. *Ann Clin Transl Neurol*. 2016 Apr 27;3(6):443–54.
144. Corsiero E, Sutcliffe N, Pitzalis C, Bombardieri M. Accumulation of Self-Reactive Naïve and Memory B Cell Reveals Sequential Defects in B Cell Tolerance Checkpoints in Sjögren’s Syndrome. *PLoS ONE* [Internet]. 2014 Dec 23 [cited 2016 Sep 8];9(12). Available from: <http://www.ncbi.nlm.nih.gov/pmc/articles/PMC4275206/>
145. Yurasov S, Wardemann H, Hammersen J, Tsuiji M, Meffre E, Pascual V, et al. Defective B cell tolerance checkpoints in systemic lupus erythematosus. *J Exp Med*. 2005 Mar 7;201(5):703–11.
146. Keenan RA, Riva AD, Corleis B, Hepburn L, Licence S, Winkler TH, et al. Censoring of Autoreactive B Cell Development by the Pre-B Cell Receptor. *Science*. 2008 Aug 1;321(5889):696–9.
147. Pewzner-Jung Y, Friedmann D, Sonoda E, Jung S, Rajewsky K, Eilat D. B cell deletion, anergy, and receptor editing in “knock in” mice targeted with a germline-encoded or somatically mutated anti-DNA heavy chain. *J Immunol Baltim Md 1950*. 1998 Nov 1;161(9):4634–45.
148. Hippen KL, Schram BR, Tze LE, Pape KA, Jenkins MK, Behrens TW. In vivo assessment of the relative contributions of deletion, anergy, and editing to B cell self-tolerance. *J Immunol Baltim Md 1950*. 2005 Jul 15;175(2):909–16.
149. Pelanda R, Torres RM. Central B-Cell Tolerance: Where Selection Begins. *Cold Spring Harb Perspect Biol*. 2012 Apr 1;4(4):a007146.
150. Luning Prak ET, Monestier M, Eisenberg RA. B cell receptor editing in tolerance and autoimmunity. *Ann N Y Acad Sci*. 2011 Jan;1217:96–121.
151. Hieter PA, Korsmeyer SJ, Waldmann TA, Leder P. Human immunoglobulin kappa light-chain genes are deleted or rearranged in lambda-producing B cells. *Nature*. 1981 Apr 2;290(5805):368–72.
152. Korsmeyer SJ, Hieter PA, Sharrow SO, Goldman CK, Leder P, Waldmann TA. Normal human B cells display ordered light chain gene rearrangements and deletions. *J Exp Med*. 1982 Oct 1;156(4):975–85.
153. Langerak AW, Nadel B, De Torbal A, Wolvers-Tettero ILM, van Gastel-Mol EJ, Verhaaf B, et al. Unraveling the consecutive recombination events in the human IGK locus. *J Immunol Baltim Md 1950*. 2004 Sep 15;173(6):3878–88.

154. Tiegs SL, Russell DM, Nemazee D. Receptor editing in self-reactive bone marrow B cells. *J Exp Med*. 1993 Apr 1;177(4):1009–20.
155. Darlow JM, Stott DI. VH replacement in rearranged immunoglobulin genes. *Immunology*. 2005 Feb;114(2):155–65.
156. Lang J, Ota T, Kelly M, Strauch P, Freed BM, Torres RM, et al. Receptor editing and genetic variability in human autoreactive B cells. *J Exp Med*. 2016 Jan 11;213(1):93–108.
157. Zhang Z, Wang Y-H, Zemlin M, Findley HW, Bridges SL, Burrows PD, et al. Molecular mechanism of serial VH gene replacement. *Ann N Y Acad Sci*. 2003 Apr;987:270–3.
158. Sun A, Novobrantseva TI, Coffre M, Hewitt SL, Jensen K, Skok JA, et al. VH replacement in primary immunoglobulin repertoire diversification. *Proc Natl Acad Sci*. 2015 Feb 3;112(5):E458–66.
159. Lange MD, Huang L, Yu Y, Li S, Liao H, Zemlin M, et al. Accumulation of VH Replacement Products in IgH Genes Derived from Autoimmune Diseases and Anti-Viral Responses in Human. *Front Immunol* [Internet]. 2014 Jul 22 [cited 2016 Sep 8];5. Available from: <http://www.ncbi.nlm.nih.gov/pmc/articles/PMC4105631/>
160. Huang L, Lange MD, Zhang Z. VH Replacement Footprint Analyzer-I, a Java-Based Computer Program for Analyses of Immunoglobulin Heavy Chain Genes and Potential VH Replacement Products in Human and Mouse. *Front Immunol* [Internet]. 2014 Feb 10 [cited 2016 Sep 8];5. Available from: <http://www.ncbi.nlm.nih.gov/pmc/articles/PMC3918983/>
161. Meng W, Jayaraman S, Zhang B, Schwartz GW, Daber RD, Hershberg U, et al. Trials and Tribulations with VH Replacement. *Front Immunol* [Internet]. 2014 Jan 30 [cited 2016 Sep 8];5. Available from: <http://www.ncbi.nlm.nih.gov/pmc/articles/PMC3906580/>
162. Nemazee DA, Bürki K. Clonal deletion of B lymphocytes in a transgenic mouse bearing anti-MHC class I antibody genes. *Nature*. 1989 Feb 9;337(6207):562–6.
163. Hartley SB, Cooke MP, Fulcher DA, Harris AW, Cory S, Basten A, et al. Elimination of self-reactive B lymphocytes proceeds in two stages: arrested development and cell death. *Cell*. 1993 Feb 12;72(3):325–35.
164. Hartley SB, Crosbie J, Brink R, Kantor AB, Basten A, Goodnow CC. Elimination from peripheral lymphoid tissues of self-reactive B lymphocytes recognizing membrane-bound antigens. *Nature*. 1991 Oct 24;353(6346):765–9.
165. Cambier JC, Gauld SB, Merrell KT, Vilen BJ. B-cell anergy: from transgenic models to naturally occurring anergic B cells? *Nat Rev Immunol*. 2007 Aug;7(8):633–43.

166. Borrero M, Clarke SH. Low-affinity anti-Smith antigen B cells are regulated by anergy as opposed to developmental arrest or differentiation to B-1. *J Immunol Baltim Md 1950*. 2002 Jan 1;168(1):13–21.
167. Goodnow CC, Crosbie J, Adelstein S, Lavoie TB, Smith-Gill SJ, Brink RA, et al. Altered immunoglobulin expression and functional silencing of self-reactive B lymphocytes in transgenic mice. *Nature*. 1988 Aug 25;334(6184):676–82.
168. Gauld SB, Benschop RJ, Merrell KT, Cambier JC. Maintenance of B cell anergy requires constant antigen receptor occupancy and signaling. *Nat Immunol*. 2005 Nov;6(11):1160–7.
169. Quách TD, Manjarrez-Orduño N, Adlowitz DG, Silver L, Yang H, Wei C, et al. Anergic Responses Characterize a Large Fraction of Human Autoreactive Naive B Cells Expressing Low Levels of Surface IgM. *J Immunol*. 2011 Apr 15;186(8):4640–8.
170. Cyster JG, Goodnow CC. Antigen-induced exclusion from follicles and anergy are separate and complementary processes that influence peripheral B cell fate. *Immunity*. 1995 Dec;3(6):691–701.
171. Ekland EH, Forster R, Lipp M, Cyster JG. Requirements for follicular exclusion and competitive elimination of autoantigen-binding B cells. *J Immunol Baltim Md 1950*. 2004 Apr 15;172(8):4700–8.
172. Cyster JG, Hartley SB, Goodnow CC. Competition for follicular niches excludes self-reactive cells from the recirculating B-cell repertoire. *Nature*. 1994 Sep 29;371(6496):389–95.
173. Getahun A, Beavers NA, Larson SR, Shlomchik MJ, Cambier JC. Continuous inhibitory signaling by both SHP-1 and SHIP-1 pathways is required to maintain unresponsiveness of anergic B cells. *J Exp Med*. 2016 May 2;213(5):751–69.
174. Pao LI, Lam K-P, Henderson JM, Kutok JL, Alimzhanov M, Nitschke L, et al. B Cell-Specific Deletion of Protein-Tyrosine Phosphatase Shp1 Promotes B-1a Cell Development and Causes Systemic Autoimmunity. *Immunity*. 2007 Jul 27;27(1):35–48.
175. Maxwell MJ, Duan M, Armes JE, Anderson GP, Tarlinton DM, Hibbs ML. Genetic Segregation of Inflammatory Lung Disease and Autoimmune Disease Severity in SHIP-1<sup>-/-</sup> Mice. *J Immunol*. 2011 Jun 15;186(12):7164–75.
176. Vinuesa CG, Tangye SG, Moser B, Mackay CR. Follicular B helper T cells in antibody responses and autoimmunity. *Nat Rev Immunol*. 2005 Nov;5(11):853–65.
177. MacLennan I. Germinal-Centers. *Annu Rev Immunol*. 1994;12:117–39.

178. Koopman G, Keehnen RM, Lindhout E, Zhou DF, de Groot C, Pals ST. Germinal center B cells rescued from apoptosis by CD40 ligation or attachment to follicular dendritic cells, but not by engagement of surface immunoglobulin or adhesion receptors, become resistant to CD95-induced apoptosis. *Eur J Immunol*. 1997 Jan;27(1):1–7.
179. Nurieva RI, Liu X, Dong C. Molecular mechanisms of T-cell tolerance. *Immunol Rev*. 2011 May 1;241(1):133–44.
180. Xing Y, Hogquist KA. T-Cell Tolerance: Central and Peripheral. *Cold Spring Harb Perspect Biol*. 2012 Jun 1;4(6):a006957.
181. Tiller T, Tsuiji M, Yurasov S, Velinzon K, Nussenzweig MC, Wardemann H. Autoreactivity in human IgG+ memory B cells. *Immunity*. 2007 Feb;26(2):205–13.
182. Schroeder K, Herrmann M, Winkler TH. The role of somatic hypermutation in the generation of pathogenic antibodies in SLE. *Autoimmunity*. 2013 Mar;46(2):121–7.
183. Zhang J, Jacobi AM, Wang T, Diamond B. Pathogenic autoantibodies in systemic lupus erythematosus are derived from both self-reactive and non-self-reactive B cells. *Mol Med Camb Mass*. 2008 Dec;14(11–12):675–81.
184. Guo W, Smith D, Aviszus K, Detanico T, Heiser RA, Wysocki LJ. Somatic hypermutation as a generator of antinuclear antibodies in a murine model of systemic autoimmunity. *J Exp Med*. 2010 Sep 27;207(10):2225–37.
185. Li Y, Takahashi Y, Fujii S, Zhou Y, Hong R, Suzuki A, et al. EAF2 mediates germinal centre B-cell apoptosis to suppress excessive immune responses and prevent autoimmunity. *Nat Commun*. 2016;7:10836.
186. Russell DM, Dembić Z, Morahan G, Miller JF a. P, Bürki K, Nemazee D. Peripheral deletion of self-reactive B cells. *Nature*. 1991 Nov 28;354(6351):308–11.
187. Ota T, Ota M, Duong BH, Gavin AL, Nemazee D. Liver-expressed I $\mu$ g $\kappa$  superantigen induces tolerance of polyclonal B cells by clonal deletion not  $\kappa$  to  $\lambda$  receptor editing. *J Exp Med*. 2011 Mar 14;208(3):617–29.
188. Li Y, Ma L, Shen J, Chong AS. Peripheral deletion of mature alloreactive B cells induced by costimulation blockade. *Proc Natl Acad Sci*. 2007 Jul 17;104(29):12093–8.
189. Rice JS, Newman J, Wang C, Michael DJ, Diamond B. Receptor editing in peripheral B cell tolerance. *Proc Natl Acad Sci U S A*. 2005 Feb 1;102(5):1608–13.
190. Han S, Zheng B, Schatz DG, Spanopoulou E, Kelsoe G. Neoteny in Lymphocytes: Rag1 and Rag2 Expression in Germinal Center B Cells. *Science*. 1996 Dec 20;274(5295):2094–7.

191. Brard F, Shannon M, Prak EL, Litwin S, Weigert M. Somatic Mutation and Light Chain Rearrangement Generate Autoimmunity in Anti-Single-Stranded DNA Transgenic Mrl/lpr Mice. *J Exp Med*. 1999 Sep 6;190(5):691–704.
192. Reed JH, Jackson J, Christ D, Goodnow CC. Clonal redemption of autoantibodies by somatic hypermutation away from self-reactivity during human immunization. *J Exp Med*. 2016 Jun 27;213(7):1255–65.
193. Sabouri Z, Schofield P, Horikawa K, Spierings E, Kipling D, Randall KL, et al. Redemption of autoantibodies on anergic B cells by variable-region glycosylation and mutation away from self-reactivity. *Proc Natl Acad Sci*. 2014 Jun 24;111(25):E2567–75.
194. Joly P, Litrowski N. Pemphigus group (vulgaris, vegetans, foliaceus, herpetiformis, brasiliensis). *Clin Dermatol*. 2011 Jul;29(4):432–6.
195. Mihai S, Sitaru C. Immunopathology and molecular diagnosis of autoimmune bullous diseases. *J Cell Mol Med*. 2007 Jun;11(3):462–81.
196. Kneisel A, Hertl M. Autoimmune bullous skin diseases. Part 1: Clinical manifestations. *JDDG J Dtsch Dermatol Ges*. 2011 Oct 1;9(10):844–57.
197. Asilian A, Yoosefi A, Faghini G. Pemphigus vulgaris in Iran: epidemiology and clinical profile. *Skinmed*. 2006 Apr;5(2):69–71.
198. Salmanpour R, Shahkar H, Namazi MR, Rahman-Shenas MR. Epidemiology of pemphigus in south-western Iran: a 10-year retrospective study (1991-2000). *Int J Dermatol*. 2006 Feb;45(2):103–5.
199. Langan SM, Smeeth L, Hubbard R, Fleming KM, Smith CJP, West J. Bullous pemphigoid and pemphigus vulgaris—incidence and mortality in the UK: population based cohort study. *BMJ*. 2008 Jul 9;337:a180.
200. Marazza G, Pham HC, Schärer L, Pedrazzetti PP, Hunziker T, Trüeb RM, et al. Incidence of bullous pemphigoid and pemphigus in Switzerland: a 2-year prospective study. *Br J Dermatol*. 2009 Oct;161(4):861–8.
201. Bastuji-Garin S, Souissi R, Blum L, Turki H, Nouria R, Jomaa B, et al. Comparative epidemiology of pemphigus in Tunisia and France: unusual incidence of pemphigus foliaceus in young Tunisian women. *J Invest Dermatol*. 1995 Feb;104(2):302–5.
202. Hietanen J, Salo OP. Pemphigus: an epidemiological study of patients treated in Finnish hospitals between 1969 and 1978. *Acta Derm Venereol*. 1982;62(6):491–6.
203. Hahn-Ristic K, Rzany B, Amagai M, Bröcker E-B, Zillikens D. Increased incidence of pemphigus vulgaris in southern Europeans living in Germany compared with native Germans. *J Eur Acad Dermatol Venereol*. 2002 Jan 1;16(1):68–71.

204. Ding X, Aoki V, Mascaro JM, Lopez-Swidorski A, Diaz LA, Fairley JA. Mucosal and mucocutaneous (generalized) pemphigus vulgaris show distinct autoantibody profiles. *J Invest Dermatol.* 1997 Oct;109(4):592–6.
205. Amagai M, Tsunoda K, Zillikens D, Nagai T, Nishikawa T. The clinical phenotype of pemphigus is defined by the anti-desmoglein autoantibody profile. *J Am Acad Dermatol.* 1999 Feb;40(2 Pt 1):167–70.
206. Anhalt GJ, Labib RS, Voorhees JJ, Beals TF, Diaz LA. Induction of Pemphigus in Neonatal Mice by Passive Transfer of IgG from Patients with the Disease. *N Engl J Med.* 1982 May 20;306(20):1189–96.
207. Amagai M, Hashimoto T, Shimizu N, Nishikawa T. Absorption of pathogenic autoantibodies by the extracellular domain of pemphigus vulgaris antigen (Dsg3) produced by baculovirus. *J Clin Invest.* 1994 Jul;94(1):59–67.
208. Payne AS, Ishii K, Kacir S, Lin C, Li H, Hanakawa Y, et al. Genetic and functional characterization of human pemphigus vulgaris monoclonal autoantibodies isolated by phage display. *J Clin Invest.* 2005 Apr;115(4):888–99.
209. Yamagami J, Payne AS, Kacir S, Ishii K, Siegel DL, Stanley JR. Homologous regions of autoantibody heavy chain complementarity-determining region 3 (H-CDR3) in patients with pemphigus cause pathogenicity. *J Clin Invest.* 2010 Nov;120(11):4111–7.
210. Beutner EH, Jordon RE. DEMONSTRATION OF SKIN ANTIBODIES IN SERA OF PEMPHIGUS VULGARIS PATIENTS BY INDIRECT IMMUNOFLUORESCENT STAINING. *Proc Soc Exp Biol Med Soc Exp Biol Med N Y N.* 1964 Nov;117:505–10.
211. Beutner EH, Jordon RE, Chorzelski TP. The Immunopathology of Pemphigus and Bullous Pemphigoid\*\*From the Department of Microbiology, SUNY at Buffalo School of Medicine, Buffalo, New York, Mayo Graduate School of Medicine, Rochester, Minn., and the Department of Dermatology, Warsaw Medical School, Warsaw, Poland. *J Invest Dermatol.* 1968 Aug 1;51(2):63–80.
212. Amagai M, Klaus-Kovtun V, Stanley JR. Autoantibodies against a novel epithelial cadherin in pemphigus vulgaris, a disease of cell adhesion. *Cell.* 1991 Nov 29;67(5):869–77.
213. Rs B, Ai M. Structure and interactions of desmosomal and other cadherins. *Semin Cell Biol.* 1992 1992;3(3):157–67.
214. Nie Z, Merritt A, Rouhi-Parkouhi M, Taberner L, Garrod D. Membrane-impermeable cross-linking provides evidence for homophilic, isoform-specific binding of desmosomal cadherins in epithelial cells. *J Biol Chem.* 2011 Jan 21;286(3):2143–54.



215. Spindler V, Rötzer V, Dehner C, Kempf B, Gliem M, Radeva M, et al. Peptide-mediated desmoglein 3 crosslinking prevents pemphigus vulgaris autoantibody-induced skin blistering. *J Clin Invest*. 2013 Feb 1;123(2):800–11.
216. Kitajima Y. New insights into desmosome regulation and pemphigus blistering as a desmosome-remodeling disease. *Kaohsiung J Med Sci*. 2013 Jan;29(1):1–13.
217. Kowalczyk AP, Green KJ. Structure, Function and Regulation of Desmosomes. *Prog Mol Biol Transl Sci*. 2013;116:95–118.
218. Hameed A, Khan AA. Microscopic Nikolsky's sign. *Clin Exp Dermatol*. 1999 Jul;24(4):312–4.
219. Al-Amoudi A, Díez DC, Betts MJ, Frangakis AS. The molecular architecture of cadherins in native epidermal desmosomes. *Nature*. 2007 Dec 6;450(7171):832–7.
220. Harrison OJ, Brasch J, Lasso G, Katsamba PS, Ahlsen G, Honig B, et al. Structural basis of adhesive binding by desmocollins and desmogleins. *Proc Natl Acad Sci*. 2016 Jun 28;113(26):7160–5.
221. Mao X, Choi EJ, Payne AS. Disruption of Desmosome Assembly by Monovalent Human Pemphigus Vulgaris Monoclonal Antibodies. *J Invest Dermatol*. 2009 Apr;129(4):908–18.
222. Sato M, Aoyama Y, Kitajima Y. Assembly pathway of desmoglein 3 to desmosomes and its perturbation by pemphigus vulgaris-IgG in cultured keratinocytes, as revealed by time-lapsed labeling immunoelectron microscopy. *Lab Invest J Tech Methods Pathol*. 2000 Oct;80(10):1583–92.
223. Jolly PS, Berkowitz P, Bektas M, Lee H-E, Chua M, Diaz LA, et al. p38MAPK signaling and desmoglein-3 internalization are linked events in pemphigus acantholysis. *J Biol Chem*. 2010 Mar 19;285(12):8936–41.
224. Amagai M, Karpati S, Prussick R, Klaus-Kovtun V, Stanley JR. Autoantibodies against the amino-terminal cadherin-like binding domain of pemphigus vulgaris antigen are pathogenic. *J Clin Invest*. 1992 Sep;90(3):919–26.
225. Sekiguchi M, Futei Y, Fujii Y, Iwasaki T, Nishikawa T, Amagai M. Dominant autoimmune epitopes recognized by pemphigus antibodies map to the N-terminal adhesive region of desmogleins. *J Immunol Baltim Md 1950*. 2001 Nov 1;167(9):5439–48.
226. Ohyama B, Nishifuji K, Chan PT, Kawaguchi A, Yamashita T, Ishii N, et al. Epitope Spreading Is Rarely Found in Pemphigus Vulgaris by Large-Scale Longitudinal Study Using Desmoglein 2–Based Swapped Molecules. *J Invest Dermatol*. 2012 Apr;132(4):1158–68.

227. Kawasaki H, Tsunoda K, Hata T, Ishii K, Yamada T, Amagai M. Synergistic pathogenic effects of combined mouse monoclonal anti-desmoglein 3 IgG antibodies on pemphigus vulgaris blister formation. *J Invest Dermatol*. 2006 Dec;126(12):2621–30.
228. Kumar B, Arora S, Kumaran MS, Jain R, Dogra S. Study of desmoglein 1 and 3 antibody levels in relation to disease severity in Indian patients with pemphigus. *Indian J Dermatol Venereol Leprol*. 2006 Jun;72(3):203–6.
229. Abasq C, Mouquet H, Gilbert D, et al. ELisa testing of anti-desmoglein 1 and 3 antibodies in the management of pemphigus. *Arch Dermatol*. 2009 May 1;145(5):529–35.
230. Shimizu H, Masunaga T, Ishiko A, Kikuchi A, Hashimoto T, Nishikawa T. Pemphigus Vulgaris and Pemphigus Foliaceus Sera Show an Inversely Graded Binding Pattern to Extracellular Regions of Desmosomes in Different Layers of Human Epidermis. *J Invest Dermatol*. 1995 Aug 1;105(2):153–9.
231. Amagai M, Stanley JR. Desmoglein as a target in skin disease and beyond. *J Invest Dermatol*. 2012 Mar;132(3 0 2):776–84.
232. Amagai M, Koch PJ, Nishikawa T, Stanley JR. Pemphigus vulgaris antigen (desmoglein 3) is localized in the lower epidermis, the site of blister formation in patients. *J Invest Dermatol*. 1996 Feb;106(2):351–5.
233. Shirakata Y, Amagai M, Hanakawa Y, Nishikawa T, Hashimoto K. Lack of mucosal involvement in pemphigus foliaceus may be due to low expression of desmoglein 1. *J Invest Dermatol*. 1998 Jan;110(1):76–8.
234. Amagai M. Autoimmunity against desmosomal cadherins in pemphigus. *J Dermatol Sci*. 1999 Jun;20(2):92–102.
235. Mahoney MG, Wang Z, Rothenberger K, Koch PJ, Amagai M, Stanley JR. Explanations for the clinical and microscopic localization of lesions in pemphigus foliaceus and vulgaris. *J Clin Invest*. 1999 Feb;103(4):461–8.
236. Wu H, Wang ZH, Yan A, Lyle S, Fakharzadeh S, Wahl JK, et al. Protection against Pemphigus Foliaceus by Desmoglein 3 in Neonates. *N Engl J Med*. 2000 Jul 6;343(1):31–5.
237. Hata T, Nishifuji K, Shimoda K, Sasaki T, Yamada T, Nishikawa T, et al. Transgenic rescue of desmoglein 3 null mice with desmoglein 1 to develop a syngeneic mouse model for pemphigus vulgaris. *J Dermatol Sci*. 2011 Jul;63(1):33–9.
238. Hashimoto T, Amagai M, Garrod DR, Nishikawa T. Immunofluorescence and immunoblot studies on the reactivity of pemphigus vulgaris and pemphigus foliaceus sera with desmoglein 3 and desmoglein 1. *Epithelial Cell Biol*. 1995;4(2):63–9.

239. Plomp JJ, Huijbers MG, van der Maarel SM, Verschuuren JJ. Pathogenic IgG4 subclass autoantibodies in MuSK myasthenia gravis. *Ann N Y Acad Sci.* 2012 Dec;1275:114–22.
240. Huijbers MG, Zhang W, Klooster R, Niks EH, Friese MB, Straasheijm KR, et al. MuSK IgG4 autoantibodies cause myasthenia gravis by inhibiting binding between MuSK and Lrp4. *Proc Natl Acad Sci U S A.* 2013 Dec 17;110(51):20783–8.
241. Ferrari S, Mudde GC, Rieger M, Veyradier A, Kremer Hovinga JA, Scheiflinger F. IgG subclass distribution of anti-ADAMTS13 antibodies in patients with acquired thrombotic thrombocytopenic purpura. *J Thromb Haemost JTH.* 2009 Oct;7(10):1703–10.
242. Huijbers MG, Querol LA, Niks EH, Plomp JJ, van der Maarel SM, Graus F, et al. The expanding field of IgG4-mediated neurological autoimmune disorders. *Eur J Neurol.* 2015 Aug 1;22(8):1151–61.
243. Futei Y, Amagai M, Ishii K, Kuroda-Kinoshita K, Ohya K, Nishikawa T. Predominant IgG4 subclass in autoantibodies of pemphigus vulgaris and foliaceus. *J Dermatol Sci.* 2001 May;26(1):55–61.
244. Funakoshi T, Lunardon L, Ellebrecht CT, Nagler AR, O’Leary CE, Payne AS. Enrichment of total serum IgG4 in patients with pemphigus. *Br J Dermatol.* 2012 Dec;167(6):1245–53.
245. Bhol K, Mohimen A, Ahmed AR. Correlation of subclasses of IgG with disease activity in pemphigus vulgaris. *Dermatol Basel Switz.* 1994;189 Suppl 1:85–9.
246. Torzecka JD, Woźniak K, Kowalewski C, Waszczykowska E, Sysa-Jedrzejowska A, Pas HH, et al. Circulating pemphigus autoantibodies in healthy relatives of pemphigus patients: coincidental phenomenon with a risk of disease development? *Arch Dermatol Res.* 2007 Aug;299(5–6):239–43.
247. Green MG, Bystryń J-C. Effect of intravenous immunoglobulin therapy on serum levels of IgG1 and IgG4 antidesmoglein 1 and antidesmoglein 3 antibodies in pemphigus vulgaris. *Arch Dermatol.* 2008 Dec;144(12):1621–4.
248. Varga E-M, Kausar F, Aberer W, Zach M, Eber E, Durham SR, et al. Tolerant beekeepers display venom-specific functional IgG4 antibodies in the absence of specific IgE. *J Allergy Clin Immunol.* 2013 May;131(5):1419–21.
249. Adjobimey T, Hoerauf A. Induction of immunoglobulin G4 in human filariasis: an indicator of immunoregulation. *Ann Trop Med Parasitol.* 2010 Sep;104(6):455–64.
250. Ljungström I, Hammarström L, Kociecka W, Smith CI. The sequential appearance of IgG subclasses and IgE during the course of *Trichinella spiralis* infection. *Clin Exp Immunol.* 1988 Nov;74(2):230–5.

251. Jutel M, Jaeger L, Suck R, Meyer H, Fiebig H, Cromwell O. Allergen-specific immunotherapy with recombinant grass pollen allergens. *J Allergy Clin Immunol*. 2005 Sep;116(3):608–13.
252. Hussain R, Poindexter RW, Ottesen EA. Control of allergic reactivity in human filariasis. Predominant localization of blocking antibody to the IgG4 subclass. *J Immunol Baltim Md 1950*. 1992 May 1;148(9):2731–7.
253. Aalberse RC, Schuurman J. IgG4 breaking the rules. *Immunology*. 2002 Jan;105(1):9–19.
254. van der Zee JS, van Swieten P, Aalberse RC. Inhibition of complement activation by IgG4 antibodies. *Clin Exp Immunol*. 1986 May;64(2):415–22.
255. James LK, Till SJ. Potential Mechanisms for IgG4 Inhibition of Immediate Hypersensitivity Reactions. *Curr Allergy Asthma Rep [Internet]*. 2016 [cited 2016 Sep 19];16. Available from: <http://www.ncbi.nlm.nih.gov/pmc/articles/PMC4759210/>
256. Aalberse RC, Gaag R van der, Leeuwen J van. Serologic aspects of IgG4 antibodies. I. Prolonged immunization results in an IgG4-restricted response. *J Immunol*. 1983 Feb 1;130(2):722–6.
257. Zenzo GD, Zambruno G, Borradori L. Endemic Pemphigus Foliaceus: Towards Understanding Autoimmune Mechanisms of Disease Development. *J Invest Dermatol*. 2012 Nov;132(11):2499–502.
258. Mentink LF, de Jong MCJM, Kloosterhuis GJ, Zuiderveen J, Jonkman MF, Pas HH. Coexistence of IgA antibodies to desmogleins 1 and 3 in pemphigus vulgaris, pemphigus foliaceus and paraneoplastic pemphigus. *Br J Dermatol*. 2007 Apr;156(4):635–41.
259. Spaeth S, Riechers R, Borradori L, Zillikens D, Büdinger L, Hertl M. IgG, IgA and IgE autoantibodies against the ectodomain of desmoglein 3 in active pemphigus vulgaris. *Br J Dermatol*. 2001 Jun;144(6):1183–8.
260. Mantis NJ, Rol N, Corthésy B. Secretory IgA's complex roles in immunity and mucosal homeostasis in the gut. *Mucosal Immunol*. 2011 Nov;4(6):603–11.
261. Macpherson AJ, Uhr T. Induction of protective IgA by intestinal dendritic cells carrying commensal bacteria. *Science*. 2004 Mar 12;303(5664):1662–5.
262. Cerutti A, Rescigno M. The Biology of Intestinal Immunoglobulin A Responses. *Immunity*. 2008 Jun;28(6):740–50.
263. Berth M, Delanghe J, Langlois M, Buyzere MD. Reference Values of Serum IgA Subclasses in Caucasian Adults by Immunonephelometry. *Clin Chem*. 1999 Feb 1;45(2):309–10.

264. Cerutti A. The regulation of IgA class switching. *Nat Rev Immunol*. 2008 Jun;8(6):421–34.
265. Benckert J, Schmolka N, Kreschel C, Zoller MJ, Sturm A, Wiedenmann B, et al. The majority of intestinal IgA+ and IgG+ plasmablasts in the human gut are antigen-specific. *J Clin Invest*. 2011 May 2;121(5):1946–55.
266. Mercer TR, Clark MB, Crawford J, Brunck ME, Gerhardt DJ, Taft RJ, et al. Targeted sequencing for gene discovery and quantification using RNA CaptureSeq. *Nat Protoc*. 2014 May;9(5):989–1009.
267. Lorenzo-Redondo R, Fryer HR, Bedford T, Kim E-Y, Archer J, Kosakovsky Pond SL, et al. Persistent HIV-1 replication maintains the tissue reservoir during therapy. *Nature*. 2016 Feb 4;530(7588):51–6.
268. Gerlinger M, Horswell S, Larkin J, Rowan AJ, Salm MP, Varela I, et al. Genomic architecture and evolution of clear cell renal cell carcinomas defined by multiregion sequencing. *Nat Genet*. 2014 Mar;46(3):225–33.
269. Campbell PJ, Pleasance ED, Stephens PJ, Dicks E, Rance R, Goodhead I, et al. Subclonal phylogenetic structures in cancer revealed by ultra-deep sequencing. *Proc Natl Acad Sci*. 2008 Sep 2;105(35):13081–6.
270. Mercer TR, Gerhardt DJ, Dinger ME, Crawford J, Trapnell C, Jeddloh JA, et al. Targeted RNA sequencing reveals the deep complexity of the human transcriptome. *Nat Biotechnol*. 2011 Nov 13;30(1):99–104.
271. Arnaout R, Lee W, Cahill P, Honan T, Sparrow T, Weiland M, et al. High-Resolution Description of Antibody Heavy-Chain Repertoires in Humans. *PLoS ONE* [Internet]. 2011 Aug 4 [cited 2016 Sep 16];6(8). Available from: <http://www.ncbi.nlm.nih.gov/pmc/articles/PMC3150326/>
272. Wu Y-C, Kipling D, Leong HS, Martin V, Ademokun AA, Dunn-Walters DK. High-throughput immunoglobulin repertoire analysis distinguishes between human IgM memory and switched memory B-cell populations. *Blood*. 2010 Aug 19;116(7):1070–8.
273. DeWitt WS, Lindau P, Snyder TM, Sherwood AM, Vignali M, Carlson CS, et al. A Public Database of Memory and Naive B-Cell Receptor Sequences. *PLoS ONE* [Internet]. 2016 Aug 11 [cited 2016 Sep 18];11(8). Available from: <http://www.ncbi.nlm.nih.gov/pmc/articles/PMC4981401/>
274. Martin V, Wu Y-C (Bryan), Kipling D, Dunn-Walters D. Ageing of the B-cell repertoire. *Phil Trans R Soc B*. 2015 Sep 5;370(1676):20140237.
275. Parameswaran P, Liu Y, Roskin KM, Jackson KK, Dixit VP, Lee J-Y, et al. Convergent antibody signatures in human dengue. *Cell Host Microbe*. 2013 Jun 12;13(6):691–700.

276. Zhu J, Wu X, Zhang B, McKee K, O'Dell S, Soto C, et al. De novo identification of VRC01 class HIV-1-neutralizing antibodies by next-generation sequencing of B-cell transcripts. *Proc Natl Acad Sci*. 2013 Oct 22;110(43):E4088–97.
277. Wu X, Zhou T, Zhu J, Zhang B, Georgiev I, Wang C, et al. Focused evolution of HIV-1 neutralizing antibodies revealed by structures and deep sequencing. *Science*. 2011 Sep 16;333(6049):1593–602.
278. Galson JD, Trück J, Clutterbuck EA, Fowler A, Cerundolo V, Pollard AJ, et al. B-cell repertoire dynamics after sequential hepatitis B vaccination and evidence for cross-reactive B-cell activation. *Genome Med*. 2016;8(1):68.
279. Jackson KJL, Liu Y, Roskin KM, Glanville J, Hoh RA, Seo K, et al. Human responses to influenza vaccination show seroconversion signatures and convergent antibody rearrangements. *Cell Host Microbe*. 2014 Jul 9;16(1):105–14.
280. Jiang N, He J, Weinstein JA, Penland L, Sasaki S, He X-S, et al. Lineage Structure of the Human Antibody Repertoire in Response to Influenza Vaccination. *Sci Transl Med*. 2013 Feb 6;5(171):171ra19.
281. Laserson U, Vigneault F, Gadala-Maria D, Yaari G, Uduman M, Vander Heiden JA, et al. High-resolution antibody dynamics of vaccine-induced immune responses. *Proc Natl Acad Sci U S A*. 2014 Apr 1;111(13):4928–33.
282. Wang C, Liu Y, Cavanagh MM, Le Saux S, Qi Q, Roskin KM, et al. B-cell repertoire responses to varicella-zoster vaccination in human identical twins. *Proc Natl Acad Sci U S A*. 2015 Jan 13;112(2):500–5.
283. Palanichamy A, Apeltsin L, Kuo TC, Sirota M, Wang S, Pitts SJ, et al. Immunoglobulin class-switched B cells form an active immune axis between CNS and periphery in multiple sclerosis. *Sci Transl Med*. 2014 Aug 6;6(248):248ra106.
284. Stern JNH, Yaari G, Heiden JAV, Church G, Donahue WF, Hintzen RQ, et al. B cells populating the multiple sclerosis brain mature in the draining cervical lymph nodes. *Sci Transl Med*. 2014 Aug 6;6(248):248ra107-248ra107.
285. Lu DR, Robinson WH. Street-experienced peripheral B cells traffic to the brain. *Sci Transl Med*. 2014 Aug 6;6(248):248fs31.
286. Hershberg U, Meng W, Zhang B, Haff N, St Clair EW, Cohen PL, et al. Persistence and selection of an expanded B-cell clone in the setting of rituximab therapy for Sjögren's syndrome. *Arthritis Res Ther*. 2014;16(1):R51.
287. Tipton CM, Fucile CF, Darce J, Chida A, Ichikawa T, Gregoret I, et al. Diversity, cellular origin and autoreactivity of antibody-secreting cell expansions in acute Systemic Lupus Erythematosus. *Nat Immunol*. 2015 Jul;16(7):755–65.

288. Sequencing Systems | Sequencer Comparison Table [Internet]. [cited 2016 Sep 16]. Available from: <http://www.illumina.com/systems/sequencing.html>
289. Larimore K, McCormick MW, Robins HS, Greenberg PD. Shaping of human germline IgH repertoires revealed by deep sequencing. *J Immunol Baltim Md 1950*. 2012 Sep 15;189(6):3221–30.
290. Products - GS FLX+ System : 454 Life Sciences, a Roche Company [Internet]. [cited 2016 Sep 16]. Available from: <http://454.com/products/gs-flx-system/>
291. Glanville J, Zhai W, Berka J, Telman D, Huerta G, Mehta GR, et al. Precise determination of the diversity of a combinatorial antibody library gives insight into the human immunoglobulin repertoire. *Proc Natl Acad Sci*. 2009 Dec 1;106(48):20216–21.
292. Liao H-X, Chen X, Munshaw S, Zhang R, Marshall DJ, Vandergrift N, et al. Initial antibodies binding to HIV-1 gp41 in acutely infected subjects are polyreactive and highly mutated. *J Exp Med*. 2011 Oct 24;208(11):2237–49.
293. Weinstein JA, Jiang N, White RA, Fisher DS, Quake SR. High-Throughput Sequencing of the Zebrafish Antibody Repertoire. *Science*. 2009 May 8;324(5928):807–10.
294. Jiang N, Weinstein JA, Penland L, White RA, Fisher DS, Quake SR. Determinism and stochasticity during maturation of the zebrafish antibody repertoire. *Proc Natl Acad Sci*. 2011 Mar 29;108(13):5348–53.
295. Kaplinsky J, Li A, Sun A, Coffre M, Korolov SB, Arnaout R. Antibody repertoire deep sequencing reveals antigen-independent selection in maturing B cells. *Proc Natl Acad Sci U S A*. 2014 Jun 24;111(25):E2622–9.
296. Greiff V, Miho E, Menzel U, Reddy ST. Bioinformatic and Statistical Analysis of Adaptive Immune Repertoires. *Trends Immunol*. 2015 Nov 1;36(11):738–49.
297. Vollmers C, Sit RV, Weinstein JA, Dekker CL, Quake SR. Genetic measurement of memory B-cell recall using antibody repertoire sequencing. *Proc Natl Acad Sci U S A*. 2013 Aug 13;110(33):13463–8.
298. Chapman DG. The Estimation of Biological Populations. *Ann Math Stat*. 1954 Mar;25(1):1–15.
299. Abadi F, Botha A, Altwegg R. Revisiting the Effect of Capture Heterogeneity on Survival Estimates in Capture-Mark-Recapture Studies: Does It Matter? *PLOS ONE*. 2013 Apr 30;8(4):e62636.
300. Chao A, Pan H-Y, Chiang S-C. The Petersen-Lincoln estimator and its extension to estimate the size of a shared population. *Biom J Biom Z*. 2008 Dec;50(6):957–70.

301. Mao CX, Huang R, Zhang S. Petersen estimator, Chapman adjustment, list effects, and heterogeneity. *Biometrics*. 2016 Jun 20;
302. Gotelli NJ, Colwell RK. Quantifying biodiversity: procedures and pitfalls in the measurement and comparison of species richness. *Ecol Lett*. 2001 Jul 22;4(4):379–91.
303. Chao A, Gotelli NJ, Hsieh TC, Sander EL, Ma KH, Colwell RK, et al. Rarefaction and extrapolation with Hill numbers: a framework for sampling and estimation in species diversity studies. *Ecol Monogr*. 2014 Feb 1;84(1):45–67.
304. Hsieh TC, Ma KH, Chao A. iNEXT: an R package for rarefaction and extrapolation of species diversity (Hill numbers). *Methods Ecol Evol*. 2016 Aug 1;n/a-n/a.
305. Bunge J, Böhning D, Allen H, Foster JA. Estimating population diversity with unreliable low frequency counts. *Pac Symp Biocomput Pac Symp Biocomput*. 2012;203–12.
306. Dickie IA. Insidious effects of sequencing errors on perceived diversity in molecular surveys. *New Phytol*. 2010;188(4):916–8.
307. Good IJ. The Population Frequencies of Species and the Estimation of Population Parameters. *Biometrika*. 1953 Dec 1;40(3–4):237–64.
308. Gale WA, Sampson G. Good-turing frequency estimation without tears. *J Quant Linguist*. 1995 Jan 1;2(3):217–37.
309. Chiu C-H, Chao A. Estimating and comparing microbial diversity in the presence of sequencing errors. *PeerJ [Internet]*. 2016 Feb 1 [cited 2016 Oct 9];4. Available from: <http://www.ncbi.nlm.nih.gov/pmc/articles/PMC4741086/>
310. Briney BS, Willis JR, Finn JA, McKinney BA, Jr JEC. Tissue-Specific Expressed Antibody Variable Gene Repertoires. *PLOS ONE*. 2014 Jun 23;9(6):e100839.
311. von Büdingen H-C, Kuo TC, Sirota M, van Belle CJ, Apeltsin L, Glanville J, et al. B cell exchange across the blood-brain barrier in multiple sclerosis. *J Clin Invest*. 2012 Dec 3;122(12):4533–43.
312. Mysara M, Leys N, Raes J, Monsieurs P. IPED: a highly efficient denoising tool for Illumina MiSeq Paired-end 16S rRNA gene amplicon sequencing data. *BMC Bioinformatics*. 2016;17:192.
313. Loman NJ, Misra RV, Dallman TJ, Constantinidou C, Gharbia SE, Wain J, et al. Performance comparison of benchtop high-throughput sequencing platforms. *Nat Biotechnol*. 2012 May;30(5):434–9.



314. Lee DF, Lu J, Chang S, Loparo JJ, Xie XS. Mapping DNA polymerase errors by single-molecule sequencing. *Nucleic Acids Res.* 2016 May 16;gkw436.
315. Edgar RC. Search and clustering orders of magnitude faster than BLAST. *Bioinformatics.* 2010 Oct 1;26(19):2460–1.
316. Schramm CA, Sheng Z, Zhang Z, Mascola JR, Kwong PD, Shapiro L. SONAR: A high-throughput pipeline for inferring antibody ontogenies from longitudinal sequencing of B cell transcripts. *B Cell Biol.* 2016;7:372.
317. briney/abtools [Internet]. GitHub. [cited 2016 Sep 17]. Available from: <https://github.com/briney/abtools>
318. Kuchenbecker L, Nienen M, Hecht J, Neumann AU, Babel N, Reinert K, et al. IMSEQ—a fast and error aware approach to immunogenetic sequence analysis. *Bioinformatics.* 2015 Sep 15;31(18):2963–71.
319. Bolotin DA, Poslavsky S, Mitrophanov I, Shugay M, Mamedov IZ, Putintseva EV, et al. MiXCR: software for comprehensive adaptive immunity profiling. *Nat Methods.* 2015 May;12(5):380–1.
320. Safonova Y, Bonissone S, Kurpilyansky E, Starostina E, Lapidus A, Stinson J, et al. IgRepertoireConstructor: a novel algorithm for antibody repertoire construction and immunoproteogenomics analysis. *Bioinformatics.* 2015 Jun 15;31(12):i53–61.
321. Ben-Dor A, Shamir R, Yakhini Z. Clustering gene expression patterns. *J Comput Biol J Comput Mol Cell Biol.* 1999 Fall-Winter;6(3–4):281–97.
322. Shlemov A, Bankevich S, Bzikadze A, Safonova Y. New algorithmic challenges of adaptive immune repertoire construction. *ResearchGate [Internet].* 2016 Apr 7 [cited 2016 Sep 17]; Available from: [https://www.researchgate.net/publication/301875978\\_New\\_algorithmic\\_challenges\\_of\\_adaptive\\_immune\\_repertoire\\_construction](https://www.researchgate.net/publication/301875978_New_algorithmic_challenges_of_adaptive_immune_repertoire_construction)
323. Islam S, Zeisel A, Joost S, La Manno G, Zajac P, Kasper M, et al. Quantitative single-cell RNA-seq with unique molecular identifiers. *Nat Methods.* 2014 Feb;11(2):163–6.
324. Kivioja T, Vähärautio A, Karlsson K, Bonke M, Enge M, Linnarsson S, et al. Counting absolute numbers of molecules using unique molecular identifiers. *Nat Methods.* 2012 Jan;9(1):72–4.
325. Kinde I, Wu J, Papadopoulos N, Kinzler KW, Vogelstein B. Detection and quantification of rare mutations with massively parallel sequencing. *Proc Natl Acad Sci U S A.* 2011 Jun 7;108(23):9530–5.

326. Khan TA, Friedensohn S, Vries ARG de, Straszewski J, Ruscheweyh H-J, Reddy ST. Accurate and predictive antibody repertoire profiling by molecular amplification fingerprinting. *Sci Adv*. 2016 Mar 1;2(3):e1501371.
327. Shugay M, Britanova OV, Merzlyak EM, Turchaninova MA, Mamedov IZ, Tuganbaev TR, et al. Towards error-free profiling of immune repertoires. *Nat Methods*. 2014 Jun;11(6):653–5.
328. Turchaninova MA, Davydov A, Britanova OV, Shugay M, Bikos V, Egorov ES, et al. High-quality full-length immunoglobulin profiling with unique molecular barcoding. *Nat Protoc*. 2016 Sep;11(9):1599–616.
329. Casbon JA, Osborne RJ, Brenner S, Lichtenstein CP. A method for counting PCR template molecules with application to next-generation sequencing. *Nucleic Acids Res*. 2011 Apr 13;gkr217.
330. Sheward DJ, Murrell B, Williamson C. Degenerate Primer IDs and the Birthday Problem. *Proc Natl Acad Sci*. 2012 May 22;109(21):E1330–E1330.
331. Jabara CB, Swanstrom R. Reply to Sheward et al.: Two sampling issues in using Primer IDs. *Proc Natl Acad Sci*. 2012 May 22;109(21):E1331–E1331.
332. Brodin J, Hedskog C, Heddini A, Benard E, Neher RA, Mild M, et al. Challenges with Using Primer IDs to Improve Accuracy of Next Generation Sequencing. *PLoS ONE* [Internet]. 2015 Mar 5 [cited 2016 Sep 17];10(3). Available from: <http://www.ncbi.nlm.nih.gov/pmc/articles/PMC4351057/>
333. Kou R, Lam H, Duan H, Ye L, Jongkam N, Chen W, et al. Benefits and Challenges with Applying Unique Molecular Identifiers in Next Generation Sequencing to Detect Low Frequency Mutations. *PLOS ONE*. 2016 Jan 11;11(1):e0146638.
334. Benichou J, Ben-Hamo R, Louzoun Y, Efroni S. Rep-Seq: uncovering the immunological repertoire through next-generation sequencing. *Immunology*. 2012 Mar;135(3):183–91.
335. Smith TS, Heger A, Sudbery I. UMI-tools: Modelling sequencing errors in Unique Molecular Identifiers to improve quantification accuracy. *bioRxiv*. 2016 May 9;51755.
336. Lan F, Haliburton JR, Yuan A, Abate AR. Droplet barcoding for massively parallel single-molecule deep sequencing. *Nat Commun*. 2016 Jun 29;7:11784.
337. Gadala-Maria D, Yaari G, Uduman M, Kleinstein SH. Automated analysis of high-throughput B-cell sequencing data reveals a high frequency of novel immunoglobulin V gene segment alleles. *Proc Natl Acad Sci U S A*. 2015 Feb 24;112(8):E862–70.

338. Ye J, Ma N, Madden TL, Ostell JM. IgBLAST: an immunoglobulin variable domain sequence analysis tool. *Nucleic Acids Res.* 2013 Jul;41(Web Server issue):W34-40.
339. Alamyar E, Duroux P, Lefranc M-P, Giudicelli V. IMGT® Tools for the Nucleotide Analysis of Immunoglobulin (IG) and T Cell Receptor (TR) V-(D)-J Repertoires, Polymorphisms, and IG Mutations: IMGT/V-QUEST and IMGT/HighV-QUEST for NGS. In: Christiansen FT, Tait BD, editors. *Immunogenetics* [Internet]. Humana Press; 2012 [cited 2016 Sep 17]. p. 569–604. (Methods in Molecular Biology). Available from: [http://dx.doi.org/10.1007/978-1-61779-842-9\\_32](http://dx.doi.org/10.1007/978-1-61779-842-9_32)
340. Wang X, Wu D, Zheng S, Sun J, Tao L, Li Y, et al. Ab-origin: an enhanced tool to identify the sourcing gene segments in germline for rearranged antibodies. *BMC Bioinformatics.* 2008;9(12):1–9.
341. Souto-Carneiro MM, Longo NS, Russ DE, Sun H, Lipsky PE. Characterization of the human Ig heavy chain antigen binding complementarity determining region 3 using a newly developed software algorithm, JOINSOLVER. *J Immunol Baltim Md* 1950. 2004 Jun 1;172(11):6790–802.
342. Russ DE, Ho K-Y, Longo NS. HTJoinSolver: Human immunoglobulin VDJ partitioning using approximate dynamic programming constrained by conserved motifs. *BMC Bioinformatics.* 2015;16:170.
343. Munshaw S, Kepler TB. SoDA2: a Hidden Markov Model approach for identification of immunoglobulin rearrangements. *Bioinforma Oxf Engl.* 2010 Apr 1;26(7):867–72.
344. Gaëta BA, Malming HR, Jackson KJL, Bain ME, Wilson P, Collins AM. iHMMune-align: hidden Markov model-based alignment and identification of germline genes in rearranged immunoglobulin gene sequences. *Bioinforma Oxf Engl.* 2007 Jul 1;23(13):1580–7.
345. Ralph DK, Iv FAM. Consistency of VDJ Rearrangement and Substitution Parameters Enables Accurate B Cell Receptor Sequence Annotation. *PLOS Comput Biol.* 2016 Jan 11;12(1):e1004409.
346. Chen Z, Collins AM, Wang Y, Gaëta BA. Clustering-based identification of clonally-related immunoglobulin gene sequence sets. *Immunome Res.* 2010 Sep 27;6(Suppl 1):S4.
347. Sok D, Laserson U, Laserson J, Liu Y, Vigneault F, Julien J-P, et al. The Effects of Somatic Hypermutation on Neutralization and Binding in the PGT121 Family of Broadly Neutralizing HIV Antibodies. *PLOS Pathog.* 2013 Nov 21;9(11):e1003754.
348. Frost SDW, Murrell B, Hossain ASMM, Silverman GJ, Pond SLK. Assigning and visualizing germline genes in antibody repertoires. *Philos Trans R Soc B Biol Sci*

- [Internet]. 2015 Sep 5 [cited 2016 Sep 18];370(1676). Available from: <http://www.ncbi.nlm.nih.gov/pmc/articles/PMC4528417/>
349. Briney B, Le K, Zhu J, Burton DR. Clonify: unseeded antibody lineage assignment from next-generation sequencing data. *Sci Rep*. 2016 Apr 22;6:23901.
  350. Felsenstein J. [24] Inferring phylogenies from protein sequences by parsimony, distance, and likelihood methods. In: *Enzymology B-M* in, editor. Academic Press; 1996 [cited 2016 Sep 18]. p. 418–27. (*Computer Methods for Macromolecular Sequence Analysis*; vol. 266). Available from: <http://www.sciencedirect.com/science/article/pii/S0076687996660261>
  351. Day WHE, Johnson DS, Sankoff D. The computational complexity of inferring rooted phylogenies by parsimony. *Math Biosci*. 1986 Sep 1;81(1):33–42.
  352. Kolaczkowski B, Thornton JW. Performance of maximum parsimony and likelihood phylogenetics when evolution is heterogeneous. *Nature*. 2004 Oct 21;431(7011):980–4.
  353. Wu X, Zhang Z, Schramm CA, Joyce MG, Do Kwon Y, Zhou T, et al. Maturation and diversity of the VRC01-antibody lineage over 15 years of chronic HIV-1 infection. *Cell*. 2015 Apr 23;161(3):470–85.
  354. Casadevall A, Janda A. Immunoglobulin isotype influences affinity and specificity. *Proc Natl Acad Sci*. 2012 Jul 31;109(31):12272–3.
  355. Tudor D, Yu H, Maupetit J, Drillet A-S, Bouceba T, Schwartz-Cornil I, et al. Isotype modulates epitope specificity, affinity, and antiviral activities of anti-HIV-1 human broadly neutralizing 2F5 antibody. *Proc Natl Acad Sci U S A*. 2012 Jul 31;109(31):12680–5.
  356. Xia Y, Eryilmaz E, Zhang Q, Cowburn D, Putterman C. Anti-DNA antibody mediated catalysis is isotype dependent. *Mol Immunol*. 2016 Jan;69:33–43.
  357. Xia Y, Pawar RD, Nakouzi AS, Herlitz L, Broder A, Liu K, et al. The constant region contributes to the antigenic specificity and renal pathogenicity of murine anti-DNA antibodies. *J Autoimmun*. 2012 Dec;39(4):398–411.
  358. Xia Y, Janda A, Eryilmaz E, Casadevall A, Putterman C. The constant region affects antigen binding of antibodies to DNA by altering secondary structure. *Mol Immunol*. 2013 Nov;56(1–2):28–37.
  359. Hammers CM, Stanley JR. Antibody phage display: technique and applications. *J Invest Dermatol*. 2014 Feb;134(2):e17.
  360. Ostertag EM, Kacir S, Thiboutot M, Gulendran G, Zheng XL, Cines DB, et al. ADAMTS13 autoantibodies cloned from patients with acquired thrombotic

- thrombocytopenic purpura: 1. Structural and functional characterization in vitro. *Transfusion (Paris)*. 2016 Jul;56(7):1763–74.
361. Roark JH, Bussel JB, Cines DB, Siegel DL. Genetic analysis of autoantibodies in idiopathic thrombocytopenic purpura reveals evidence of clonal expansion and somatic mutation. *Blood*. 2002 Aug 15;100(4):1388–98.
362. Ishii K, Lin C, Siegel DL, Stanley JR. Isolation of pathogenic monoclonal anti-desmoglein 1 human antibodies by phage display of pemphigus foliaceus autoantibodies. *J Invest Dermatol*. 2008 Apr;128(4):939–48.
363. Roben P, Barbas SM, Sandoval L, Lecerf JM, Stollar BD, Solomon A, et al. Repertoire cloning of lupus anti-DNA autoantibodies. *J Clin Invest*. 1996 Dec 15;98(12):2827–37.
364. del Rincon I, Zeidel M, Rey E, Harley JB, James JA, Fischbach M, et al. Delineation of the human systemic lupus erythematosus anti-Smith antibody response using phage-display combinatorial libraries. *J Immunol Baltim Md 1950*. 2000 Dec 15;165(12):7011–6.
365. Stanley JR, Ishii K, Siegel DL, Payne AS. Update on the cloning of monoclonal anti-desmoglein antibodies from human pemphigus patients: implications for targeted therapy. *Vet Dermatol*. 2009 Oct;20(5–6):327–30.
366. Payne AS. Cloning and genetic characterization of human pemphigus autoantibodies. *J Am Acad Dermatol*. 2006 Aug;55(2):e2.
367. Cho MJ, Lo ASY, Mao X, Nagler AR, Ellebrecht CT, Mukherjee EM, et al. Shared VH1-46 gene usage by pemphigus vulgaris autoantibodies indicates common humoral immune responses among patients. *Nat Commun*. 2014 Jun 19;5:4167.
368. Payne AS, Siegel DL, Stanley JR. Targeting pemphigus autoantibodies through their heavy-chain variable region genes. *J Invest Dermatol*. 2007 Jul;127(7):1681–91.
369. Noel D, Bernardi T, Navarro-Teulon I, Marin M, Martinetto J-P, Ducancel F, et al. Analysis of the individual contributions of immunoglobulin heavy and light chains to the binding of antigen using cell transfection and plasmon resonance analysis. *J Immunol Methods*. 1996 Jun 21;193(2):177–87.
370. Hoogenboom HR, Winter G. By-passing immunisation. *J Mol Biol*. 1992 Sep 20;227(2):381–8.
371. Nissim A, Hoogenboom HR, Tomlinson IM, Flynn G, Midgley C, Lane D, et al. Antibody fragments from a “single pot” phage display library as immunochemical reagents. *EMBO J*. 1994 Feb 1;13(3):692–8.
372. Xu JL, Davis MM. Diversity in the CDR3 Region of VH Is Sufficient for Most Antibody Specificities. *Immunity*. 2000 Jul 1;13(1):37–45.

373. Chen K, Xu W, Wilson M, He B, Miller NW, Bengtén E, et al. Immunoglobulin D enhances immune surveillance by activating antimicrobial, proinflammatory and B cell-stimulating programs in basophils. *Nat Immunol*. 2009 Aug;10(8):889–98.
383. Payne AS, Stanley JR. Pemphigus. In: Goldsmith L, Katz S, Gilchrest B, Paller A, Leffell D, Wolff K. *Fitzpatrick's Dermatology in General Medicine*, Eighth Edition, 2 Volume set. 8 edition. New York: McGraw-Hill Education / Medical; 2012. 2 p.
375. Bhol KC, Ahmed AR. Production of non-pathogenic human monoclonal antibodies to desmoglein 3 from pemphigus vulgaris patient. *Autoimmunity*. 2002 Mar;35(2):87–91.
376. Di Zenzo G, Di Lullo G, Corti D, Calabresi V, Sinistro A, Vanzetta F, et al. Pemphigus autoantibodies generated through somatic mutations target the desmoglein-3 cis-interface. *J Clin Invest*. 2012 Oct;122(10):3781–90.
377. Mascaró JM, España A, Liu Z, Ding X, Swartz SJ, Fairley JA, et al. Mechanisms of acantholysis in pemphigus vulgaris: role of IgG valence. *Clin Immunol Immunopathol*. 1997 Oct;85(1):90–6.
378. Kricheli D, David M, Frusic-Zlotkin M, Goldsmith D, Rabinov M, Sulkes J, et al. The distribution of pemphigus vulgaris-IgG subclasses and their reactivity with desmoglein 3 and 1 in pemphigus patients and their first-degree relatives. *Br J Dermatol*. 2000 Aug;143(2):337–42.
379. Yeh S-W, Cavacini LA, Bhol KC, Lin M-S, Kumar M, Duval M, et al. Pathogenic human monoclonal antibody against desmoglein 3. *Clin Immunol Orlando Fla*. 2006 Jul;120(1):68–75.
380. Liu F, Bergami PL, Duval M, Kuhrt D, Posner M, Cavacini L. Expression and functional activity of isotype and subclass switched human monoclonal antibody reactive with the base of the V3 loop of HIV-1 gp120. *AIDS Res Hum Retroviruses*. 2003 Jul;19(7):597–607.
381. Miranda LR, Duval M, Doherty H, Seaman MS, Posner MR, Cavacini LA. The neutralization properties of a HIV-specific antibody are markedly altered by glycosylation events outside the antigen-binding domain. *J Immunol Baltim Md 1950*. 2007 Jun 1;178(11):7132–8.
382. Mao X, Li H, Sano Y, Gaestel M, Mo Park J, Payne AS. MAPKAP kinase 2 (MK2)-dependent and -independent models of blister formation in pemphigus vulgaris. *J Invest Dermatol*. 2014 Jan;134(1):68–76.
383. Rader C, Popkov M, Neves JA, Barbas CF. Integrin alpha(v)beta3 targeted therapy for Kaposi's sarcoma with an in vitro evolved antibody. *FASEB J Off Publ Fed Am Soc Exp Biol*. 2002 Dec;16(14):2000–2.

384. Futei Y, Amagai M, Sekiguchi M, Nishifuji K, Fujii Y, Nishikawa T. Use of domain-swapped molecules for conformational epitope mapping of desmoglein 3 in pemphigus vulgaris. *J Invest Dermatol*. 2000 Nov;115(5):829–34.
385. Hanakawa Y, Matsuyoshi N, Stanley JR. Expression of desmoglein 1 compensates for genetic loss of desmoglein 3 in keratinocyte adhesion. *J Invest Dermatol*. 2002 Jul;119(1):27–31.
386. Amagai M, Matsuyoshi N, Wang ZH, Andl C, Stanley JR. Toxin in bullous impetigo and staphylococcal scalded-skin syndrome targets desmoglein 1. *Nat Med*. 2000 Nov;6(11):1275–7.
387. Sharma PM, Choi EJ, Kuroda K, Hachiya T, Ishii K, Payne AS. Pathogenic anti-desmoglein monoclonal antibodies demonstrate variable ELISA activity due to preferential binding of mature versus proprotein isoforms of desmoglein 3. *J Invest Dermatol*. 2009 Sep;129(9):2309–12.
388. Stavnezer J, Guikema JEJ, Schrader CE. Mechanism and Regulation of Class Switch Recombination. *Annu Rev Immunol*. 2008;26:261–92.
389. Tavakolpour S, Tavakolpour V. Interleukin 4 inhibition as a potential therapeutic in pemphigus. *Cytokine*. 2016 Jan;77:189–95.
390. Zhang K, Mills FC, Saxon A. Switch circles from IL-4-directed epsilon class switching from human B lymphocytes. Evidence for direct, sequential, and multiple step sequential switch from mu to epsilon Ig heavy chain gene. *J Immunol Baltim Md 1950*. 1994 Apr 1;152(7):3427–35.
391. Jeannin P, Lecoanet S, Delneste Y, Gauchat JF, Bonnefoy JY. IgE versus IgG4 production can be differentially regulated by IL-10. *J Immunol Baltim Md 1950*. 1998 Apr 1;160(7):3555–61.
392. Tavakolpour S. Interleukin 21 as a new possible player in pemphigus: Is it a suitable target? *Int Immunopharmacol*. 2016 May;34:139–45.
393. Hertl M, Eming R, Veldman C. T cell control in autoimmune bullous skin disorders. *J Clin Invest*. 2006 May;116(5):1159–66.
394. Takahashi H, Amagai M, Tanikawa A, Suzuki S, Ikeda Y, Nishikawa T, et al. T Helper Type 2-Biased Natural Killer Cell Phenotype in Patients with Pemphigus Vulgaris. *J Invest Dermatol*. 2007 Feb;127(2):324–30.
395. Goodnow CC, Sprent J, Fazekas de St Groth B, Vinuesa CG. Cellular and genetic mechanisms of self tolerance and autoimmunity. *Nature*. 2005 Jun 2;435(7042):590–7.
396. Schroeder Jr. HW, Cavacini L. Structure and function of immunoglobulins. *J Allergy Clin Immunol*. 2010 Feb;125(2, Supplement 2):S41–52.

397. Crisp SJ, Kullmann DM, Vincent A. Autoimmune synaptopathies. *Nat Rev Neurosci*. 2016 Feb;17(2):103–17.
398. Fujieda S, Zhang K, Saxon A. IL-4 plus CD40 monoclonal antibody induces human B cells gamma subclass-specific isotype switch: switching to gamma 1, gamma 3, and gamma 4, but not gamma 2. *J Immunol Baltim Md 1950*. 1995 Sep 1;155(5):2318–28.
399. Kracker S, Radbruch A. Immunoglobulin class switching: in vitro induction and analysis. *Methods Mol Biol Clifton NJ*. 2004;271:149–59.
400. McHeyzer-Williams LJ, McHeyzer-Williams MG. Antigen-specific memory B cell development. *Annu Rev Immunol*. 2005;23:487–513.
401. Cho MJ, Ellebrecht CT, Hammers CM, Mukherjee EM, Sapparapu G, Boudreaux CE, et al. Determinants of VH1-46 Cross-Reactivity to Pemphigus Vulgaris Autoantigen Desmoglein 3 and Rotavirus Antigen VP6. *J Immunol*. 2016 Jul 11;1600567.
402. Zhang J, Kobert K, Flouri T, Stamatakis A. PEAR: a fast and accurate Illumina Paired-End reAd mergeR. *Bioinformatics*. 2014 Mar 1;30(5):614–20.
403. Felsenstein J. PHYLIP - Phylogeny Inference Package (Version 3.2). *Cladistics*. 1989;5:164–6.
404. Qian Y, Jeong JS, Ye J, Dang B, Abdeladhim M, Aoki V, et al. Overlapping IgG4 Responses to Self- and Environmental Antigens in Endemic Pemphigus Foliaceus. *J Immunol Baltim Md 1950*. 2016 Mar 1;196(5):2041–50.
405. Aoki V, Rivitti EA, Diaz LA, the Cooperative Group on Fogo Selvagem Research. Update on fogo selvagem, an endemic form of pemphigus foliaceus. *J Dermatol*. 2015 Jan 1;42(1):18–26.
406. Warren SJP, Arteaga LA, Rivitti EA, Aoki V, Hans-Filho G, Qaqish BF, et al. The role of subclass switching in the pathogenesis of endemic pemphigus foliaceus. *J Invest Dermatol*. 2003 Jan;120(1):104–8.
407. van de Veen W, Stanic B, Yaman G, Wawrzyniak M, Söllner S, Akdis DG, et al. IgG4 production is confined to human IL-10-producing regulatory B cells that suppress antigen-specific immune responses. *J Allergy Clin Immunol*. 2013 Apr;131(4):1204–12.
408. Zhu H-Q, Xu R-C, Chen Y-Y, Yuan H-J, Cao H, Zhao X-Q, et al. Impaired function of CD19(+) CD24(hi) CD38(hi) regulatory B cells in patients with pemphigus. *Br J Dermatol*. 2015 Jan;172(1):101–10.



409. Iwata Y, Matsushita T, Horikawa M, Dilillo DJ, Yanaba K, Venturi GM, et al. Characterization of a rare IL-10-competent B-cell subset in humans that parallels mouse regulatory B10 cells. *Blood*. 2011 Jan 13;117(2):530–41.
410. Zan H, Cerutti A, Dramitinos P, Schaffer A, Casali P. CD40 Engagement Triggers Switching to IgA1 and IgA2 in Human B Cells Through Induction of Endogenous TGF- $\beta$ : Evidence for TGF- $\beta$  But Not IL-10-Dependent Direct S $\mu$ →S $\alpha$  and Sequential S $\mu$ →S $\gamma$ , S $\gamma$ →S $\alpha$  DNA Recombination. *J Immunol Baltim Md 1950*. 1998 Nov 15;161(10):5217–25.
411. Giordano CN, Sinha AA. Cytokine networks in Pemphigus vulgaris: An integrated viewpoint. *Autoimmunity*. 2012 Sep;45(6):427–39.
412. D’Auria L, Bonifati C, Mussi A, D’Agosto G, De Simone C, Giacalone B, et al. Cytokines in the sera of patients with pemphigus vulgaris: interleukin-6 and tumour necrosis factor-alpha levels are significantly increased as compared to healthy subjects and correlate with disease activity. *Eur Cytokine Netw*. 1997 Dec;8(4):383–7.
413. Asashima N, Fujimoto M, Watanabe R, Nakashima H, Yazawa N, Okochi H, et al. Serum levels of BAFF are increased in bullous pemphigoid but not in pemphigus vulgaris. *Br J Dermatol*. 2006 Aug;155(2):330–6.
414. Watanabe R, Fujimoto M, Yazawa N, Nakashima H, Asashima N, Kuwano Y, et al. Increased serum levels of a proliferation-inducing ligand in patients with bullous pemphigoid. *J Dermatol Sci*. 2007 Apr;46(1):53–60.
415. Corthésy B. Multi-Faceted Functions of Secretory IgA at Mucosal Surfaces. *Front Immunol [Internet]*. 2013 Jul 12 [cited 2016 Sep 16];4. Available from: <http://www.ncbi.nlm.nih.gov/pmc/articles/PMC3709412/>
416. Henry Dunand CJ, Wilson PC. Restricted, canonical, stereotyped and convergent immunoglobulin responses. *Philos Trans R Soc Lond B Biol Sci*. 2015 Sep 5;370(1676).
417. Strauli NB, Hernandez RD. Statistical inference of a convergent antibody repertoire response to influenza vaccine. *Genome Med*. 2016;8(1):60.
418. Scheid JF, Mouquet H, Ueberheide B, Diskin R, Klein F, Oliveira TYK, et al. Sequence and structural convergence of broad and potent HIV antibodies that mimic CD4 binding. *Science*. 2011 Sep 16;333(6049):1633–7.
419. Pugh-Bernard AE, Silverman GJ, Cappione AJ, Villano ME, Ryan DH, Insel RA, et al. Regulation of inherently autoreactive VH4-34 B cells in the maintenance of human B cell tolerance. *J Clin Invest*. 2001 Oct 1;108(7):1061–70.
420. Thompson KM, Sutherland J, Barden G, Melamed MD, Randen I, Natvig JB, et al. Human Monoclonal Antibodies against Blood Group Antigens Preferentially

Express a VH4-21 Variable Region Gene-Associated Epitope. *Scand J Immunol*. 1991 Oct 1;34(4):509–18.

421. Steinsbø Ø, Dunand CJH, Huang M, Mesin L, Salgado-Ferrer M, Lundin KEA, et al. Restricted VH/VL usage and limited mutations in gluten-specific IgA of coeliac disease lesion plasma cells. *Nat Commun* [Internet]. 2014 Jun 9 [cited 2016 Oct 15];5. Available from: <http://www.nature.com/doi/10.1038/ncomms5041>
422. Coales SJ, Tuske SJ, Tomasso JC, Hamuro Y. Epitope mapping by amide hydrogen/deuterium exchange coupled with immobilization of antibody, on-line proteolysis, liquid chromatography and mass spectrometry. *Rapid Commun Mass Spectrom RCM*. 2009 Mar;23(5):639–47.
423. Pandit D, Tuske SJ, Coales SJ, E SY, Liu A, Lee JE, et al. Mapping of discontinuous conformational epitopes by amide hydrogen/deuterium exchange mass spectrometry and computational docking. *J Mol Recognit JMR*. 2012 Mar;25(3):114–24.
424. Weitkamp J-H, Kallewaard N, Kusuhara K, Bures E, Williams JV, LaFleur B, et al. Infant and adult human B cell responses to rotavirus share common immunodominant variable gene repertoires. *J Immunol Baltim Md 1950*. 2003 Nov 1;171(9):4680–8.
425. Berkowska MA, Schickel J-N, Grosserichter-Wagener C, Ridder D de, Ng YS, Dongen JJM van, et al. Circulating Human CD27–IgA+ Memory B Cells Recognize Bacteria with Polyreactive Igs. *J Immunol*. 2015 Jul 6;1402708.
426. Quan CP, Berneman A, Pires R, Avrameas S, Bouvet JP. Natural polyreactive secretory immunoglobulin A autoantibodies as a possible barrier to infection in humans. *Infect Immun*. 1997 Oct 1;65(10):3997–4004.
427. McDaniel JR, DeKosky BJ, Tanno H, Ellington AD, Georgiou G. Ultra-high-throughput sequencing of the immune receptor repertoire from millions of lymphocytes. *Nat Protoc*. 2016 Mar;11(3):429–42.
428. Sun M, Li L, Gao QS, Paul S. Antigen recognition by an antibody light chain. *J Biol Chem*. 1994 Jan 7;269(1):734–8.
429. Song MK, Oh MS, Lee JH, Lee JN, Chung JH, Park SG, et al. Light chain of natural antibody plays a dominant role in protein antigen binding. *Biochem Biophys Res Commun*. 2000 Feb 16;268(2):390–4.
430. Bruin R de, Spelt K, Mol J, Koes R, Quattrocchio F. Selection of high-affinity phage antibodies from phage display libraries. *Nat Biotechnol*. 1999 Apr;17(4):397–9.

431. Derda R, Tang SKY, Li SC, Ng S, Matochko W, Jafari MR. Diversity of Phage-Displayed Libraries of Peptides during Panning and Amplification. *Molecules*. 2011 Feb 21;16(2):1776–803.
432. Kodituwakku AP, Jessup C, Zola H, Robertson DM. Isolation of antigen-specific B cells. *Immunol Cell Biol*. 2003 Jun;81(3):163–70.
433. Rhiel L, Krah S, Günther R, Becker S, Kolmar H, Hock B. REAL-Select: Full-Length Antibody Display and Library Screening by Surface Capture on Yeast Cells. *PLOS ONE*. 2014 Dec 12;9(12):e114887.
434. Kanamori T, Fujino Y, Ueda T. PURE ribosome display and its application in antibody technology. *Biochim Biophys Acta BBA - Proteins Proteomics*. 2014 Nov;1844(11):1925–32.
435. Zhou C, Shen WD. Mammalian cell surface display of full length IgG. *Methods Mol Biol Clifton NJ*. 2012;907:293–302.
436. Akdis CA, Akdis M. Mechanisms of allergen-specific immunotherapy and immune tolerance to allergens. *World Allergy Organ J [Internet]*. 2015 May 14 [cited 2016 Oct 16];8(1). Available from: <http://www.ncbi.nlm.nih.gov/pmc/articles/PMC4430874/>
437. Figueiredo JP, Oliveira RR, Cardoso LS, Barnes KC, Grant AV, Carvalho EM, et al. Adult worm-specific IgE/IgG4 balance is associated with low infection levels of *Schistosoma mansoni* in an endemic area. *Parasite Immunol*. 2012 Dec;34(12):604–10.
438. Akdis CA, Blesken T, Akdis M, Wüthrich B, Blaser K. Role of interleukin 10 in specific immunotherapy. *J Clin Invest*. 1998 Jul 1;102(1):98–106.
439. Qian Y, Jeong JS, Maldonado M, Valenzuela JG, Gomes R, Evangelista F, et al. Brazilian pemphigus foliaceus anti-desmoglein 1 autoantibodies cross-react with sand fly salivary LJM11 antigen. *J Immunol Baltim Md 1950*. 2012 Aug 15;189(4):1535–9.
440. Sarig O, Bercovici S, Zoller L, Goldberg I, Indelman M, Nahum S, et al. Population-specific association between a polymorphic variant in ST18, encoding a pro-apoptotic molecule, and pemphigus vulgaris. *J Invest Dermatol*. 2012 Jul;132(7):1798–805.
441. Vodo D, Sarig O, Geller S, Ben-Asher E, Olender T, Bochner R, et al. Identification of a Functional Risk Variant for Pemphigus Vulgaris in the ST18 Gene. *PLoS Genet [Internet]*. 2016 May 5 [cited 2016 Oct 16];12(5). Available from: <http://www.ncbi.nlm.nih.gov/pmc/articles/PMC4858139/>
442. Veldman C, Stauber A, Wassmuth R, Uter W, Schuler G, Hertl M. Dichotomy of autoreactive Th1 and Th2 cell responses to desmoglein 3 in patients with pemphigus

vulgaris (PV) and healthy carriers of PV-associated HLA class II alleles. *J Immunol* Baltim Md 1950. 2003 Jan 1;170(1):635-42.

443. Lundgren M, Persson U, Larsson P, Magnusson C, Smith CI, Hammarström L, et al. Interleukin 4 induces synthesis of IgE and IgG4 in human B cells. *Eur J Immunol*. 1989 Jul;19(7):1311-5.