Aus dem Bereich Theoretische Medizin und Biowissenschaften der Medizinischen Fakultät der Universität des Saarlandes, Homburg/Saar, und dem Department of Infection and Immunity des Luxembourg Institute of Health, Luxembourg

Erforschung der Einsatzmöglichkeiten von B Zellen für die Diagnose der Borreliose

---

Investigation of the possible use of B cells for the diagnosis of acute Lyme disease

Dissertation

zur Erlangung des Grades des Doktors der Naturwissenschaften

der Medizinischen Fakultät

der UNIVERSITÄT DES SAARLANDES

2017

vorgelegt von

Josiane Kirpach

Geb. am 09. März 1987 in Luxemburg

Table of Contents

Index of Figures

Index of Tables

List of Abbreviations

ASC	Antibody secreting cell
BCR	B cell receptor
CDR	Complementarity determining region
CDR3aa	CDR3 amino acid sequence
CHRH3	CDR3 of the heavy chain
CLL	Chronic lymphocytic leukemia
DMSO	Dimethylsulfoxid
EDTA	Ethylenediaminetetraacetic acid
ELF	Ectopic lymphoid follicle
FR	Framework region
HC	Healthy control or Heavy chain
HIV	Human immunodeficiency virus
IR6	Invariable region 6
LC	Light chain
MHC	Major histocompatibility complex
MID	Molecular identifier
MS	Multiple sclerosis
NC	Negative control
NGS	Next generation sequencing
Osp	Outer surface protein
PAGE	Polyacrylamide gel electrophoresis
PBMCs	Peripheral blood mononuclear cells
PC	Positive control

RNAseq	RNA sequencing using NGS
RT	Reverse transcription or room temperature
SePOP	Selective Precipitation Optimized Process desalting
sMZcs	Splenic marginal zone B cells
SN	Seronegative
SP	Seropositive
T0/T1/T2	Timepoint0/1/2
TD	T cell dependent
TI	T cell independent
TLR	Toll-like receptor
Top50/200	Top 50/200 clones of repertoire
UID	Unique molecular identifier
VDJ	Variable diversity joining gene segments
VlsE	Variable major protein-like sequence expressed
VlsE-C6	peptide representing IR6 of VlsE

ZUSAMMENFASSUNG

Die Fehldiagnose der Borreliose bleibt noch immer ein relevantes Problem. Große Auseinandersetzungen über die Definition und Ursachen des chronischen Verlaufs der Krankheit sind noch immer aktuell. Die nur geringe Anzahl an *Borrelien* in Patientenproben erschwert den direkten Nachweis des Krankheitserregers. Aus diesem Grund basieren diagnostische Tests auf indirekten Nachweismethoden. Der beste und einzig offiziell anerkannte Test zur Unterstützung der Diagnose der Borreliose bleibt der Nachweis von spezifischen Antikörpern. Obwohl heutige Versionen sehr gut optimiert wurden und ziemlich zuverlässig sind, ermöglichen es diese Tests nicht, eine akute von einer geheilten Krankheit zu unterscheiden. Wegen ihrer Spezifität für Antigene sind B Zellen sehr interessante Kandidaten für neue Biomarker. Im Gegensatz zu den Antikörpern, die noch lange nach dem Immunkontakt nachweislich bleiben können, sollten die B Zellen nur während einer akuten Infektion erhöht sein. Das Ziel dieser Studie war es zu zeigen, dass durch *Borrelien* hervorgerufene B Zellen im peripheren Blut von akuten Patienten nachweislich sind, und zu testen, inwiefern ähnliche Zellen in verschiedenen Patienten vorkommen. Unsere Resultate stimmen mit denen von anderen Studien überein und zeigen, dass im Vergleich zu zufällig gewählten Individuen eine größere Wahrscheinlichkeit besteht, überlappende B Zell Klone zwischen Patienten, die dem gleichen Immun Stimulus ausgesetzt waren, zu finden. Obwohl wir die Patienten an unterschiedlichen Zeitpunkten untersucht haben, konnten wir eine größere Überlappung zwischen unseren Patienten feststellen. Unsere Daten zeigen allerdings, dass die überlappenden Klone in sehr unterschiedlichem Ausmaß in den verschiedenen Patienten vorkommen. Die Klone, die in den einzelnen Proben dominierten, waren eher Patienten spezifisch. Dies stimmt ganz gut mit der dynamischen Natur der B Zell Immunantwort überein. Da das Auftreten von identischen B Zell Rezeptor Sequenzen in verschiedenen Patienten eher selten war, wollten wir testen, ob weniger restriktive Gruppierungskriterien es ermöglichen, Sequenzen in Kategorien nach der gleichen Epitop Reaktivität einzuteilen. Es stellte sich heraus, dass die Sequenzen wahrscheinlich zu unterschiedlich sind, als dass Sie in Gruppen mit der gleichen Epitop Reaktivität eingeteilt werden können. Um dies zu erreichen müssen ausgefeiltere bioinformatische Skripte entwickelt werden. Diese müssten Mutationen an Schlüsselstellen in der Antikörpersequenz sowie die dreidimensionale Struktur des Antikörpermoleküls mit einbeziehen. Unsere Befunde deuten darauf hin, dass die gesuchten B Zell Signaturen komplexer sind als initial

erwartet. Es wird wahrscheinlich nicht möglich sein, ein paar Schlüssel CDR3 Aminosäure Sequenzen zu isolieren. Wir erwarten uns eher eine Liste an verschiedenen *Borrelia*-spezifischen B Zell Klonen, die durch Einzelmutationen an Schlüsselstellen von Klonen anderer Reaktivität unterscheidbar sind. Wegen der Polyspezifität und der dynamischen Natur der B Zell Immunantwort wird wahrscheinlich eine Kombination von verschiedenen Klonen gleichzeitig nachgewiesen werden müssen. Das Matchen von B Zell Klonen mit ihren jeweiligen Antigenen, zusammen mit der Analyse vorkommender Mutationen und des B Zell Rezeptor Isotyps, sowie die phylogenetische Verwandtschaft der einzelnen Sequenzen dieser Klone hat das Potenzial, laufende (primäre und sekundäre) von vergangenen oder autoimmun Reaktionen zu unterscheiden. Dies könnte sehr hilfreich sein um einige der Kontroversen über den chronischen Verlauf der Borreliose aufzuklären.

ABSTRACT

Misdiagnosis of acute Lyme disease still remains a huge problem. Also many debates are going on about the definition and underlying causes of the chronic course of this disease. Due to the scarcity of *Borrelia* in patient samples, direct detection of the pathogen is challenging. Diagnostic laboratories need to work with indirect approaches to prove the presence of *Borrelia* inside patients. The best and currently only accepted test to support the diagnosis of acute Lyme disease infection is the detection of *Borrelia*-reactive antibodies. Although extensive optimization has led to the development of specific and quite reliable tests to prove the presence of these molecules, they have the limitation that they do not enable to directly distinguish an ongoing from a previous infection. Since B cells are the cells of our immune system that react most specifically to immune challenges, they are very interesting candidates to be explored for biomarker research. As opposed to the antibodies, which can stay in peripheral blood for prolonged times even after clearance of the infection, *Borrelia*-reactive B cells should only be elevated when the pathogen is present. The main goal of this study was to prove the presence of *Borrelia*-reactive B cells in peripheral blood of acute Lyme disease patients and to assess to what extent expanded clones overlap between patients. Our results are in line with other B cell repertoire studies, indicating that individuals that have been challenged with the same immune stimulus have a higher chance to present overlapping clones as compared to other randomly selected donors. Since we were not able to use precise timepoints in the case of acute disease, it was quite surprising to find an increased overlap of B cell clones in our patients. Our data indicate, that overlapping clones appear at heterogeneous levels between donors. Rather patient unique B cell clones are predominating individual repertoires. This is in accordance with the rather dynamic nature of B cell immune responses. Since the occurrence of identical B cell receptor sequences in different patients was a rare event, we explored whether sequence clustering using less stringent parameters would allow to extract B cell clones reacting towards the same epitope. Differences in sequences between clones were however too large to allow reliable grouping of sequences according to epitope-reactivity. For this purpose, more sophisticated bioinformatics clustering tools need to be developed. Those need to take mutations at key residues and three dimensional structures of antibody molecules into account. From our findings, we conclude, that B cell repertoire signatures might be more complex than initially imagined. We will probably not be able to isolate a few key CDR3 amino acid signature

candidates, but we will end up with a rather large list of *Borrelia*-reactive B cell clones that can be distinguished from B cells reactive towards other antigens by the presence of single key mutations. The polyspecific and dynamic nature of the B cell immune response will probably require to prove the presence of certain combinations of clones. Matching of B cell clones with their corresponding antigens and analysis of mutation levels, isotypes and phylogenetic relationships among sequences from these clones has the potential to allow to distinguish ongoing (primary and secondary) from past immune responses and autoimmune diseases. This might help to solve some of the controversies about post treatment Lyme disease.

1 INTRODUCTION

1.1 The B cell immune response

As opposed to T cells, which only indirectly recognize processed antigens presented on MHC molecules, B cells recognize antigens directly through their B cell receptor. In the course of the immune response, B cells even further optimize their unique B cell receptor for antigen binding. These properties make B cells highly interesting candidates to explore as new biomarkers of acute infections [4]. Their usefulness in this regard has already been demonstrated by the frequent use of antibodies to support the diagnosis of diseases [5].

1.1.1 Diversification and selection of B cell receptors during B cell development

Each B cell expresses a more or less unique B cell receptor on its surface, which allows for specific recognition of virtually any non-self structure. The B cell receptor is composed of heavy (H) and light (L) chain, each one containing three variable complementarity determining regions (CDRs) [6,7]. Those are the major sites of antigen recognition. Mature naïve B cells are generated throughout life from hematopoietic stem cells in the bone marrow [8,9]. The great diversity of their B cell receptors is generated during development by genetic recombination of variable (V), diversity (D) and joining (J) gene segments [10] (Figure 1). Imprecise joining between the segments as well as nucleotide insertions at the junctions of mainly the heavy chain further increase diversity of the receptor [10]. The CDR1 and CDR2 regions of both heavy and light chains are encoded within the V genes, while the CDR3s form the combination sites between VDJ on the heavy chain and VJ on the light chain. The CDR3 of the heavy chain (CDRH3) is the most variable part of the antibody and therefore most likely responsible for the specificity of antigen recognition [11]. For this reason many studies focus on this region to characterize and identify B cell clones of interest.

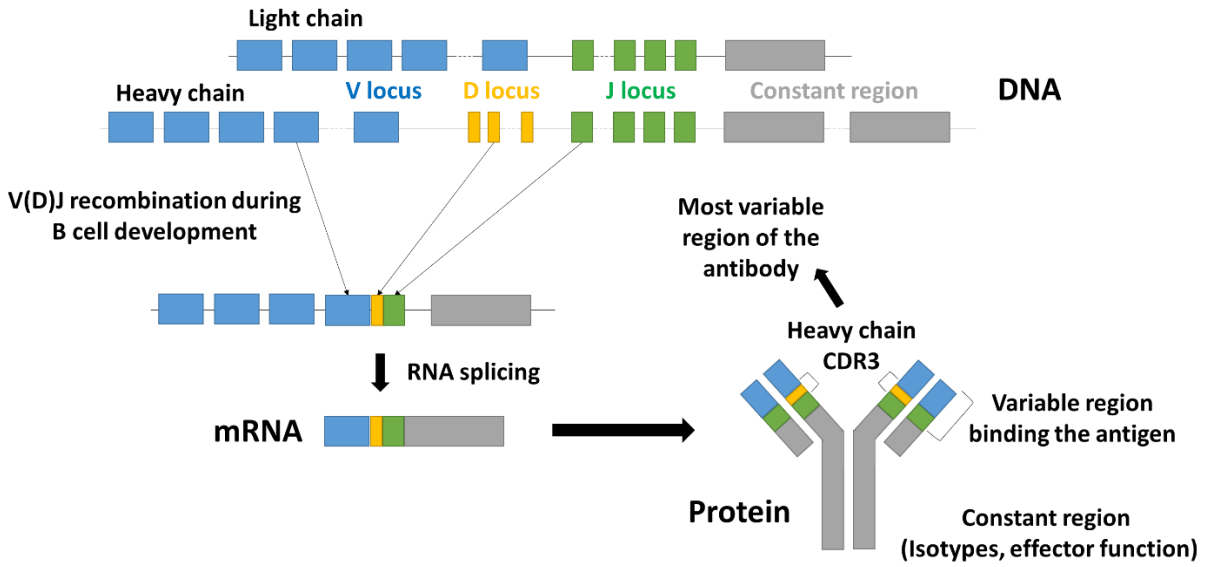


Figure 1: Recombination of gene segments during B cell development allows to generate the great diversity of B cell receptors that form the naïve repertoire.

V(D)J recombination is not a totally stochastic process, as a similar bias in gene segment usage can be observed between individuals and at different timepoints [12,13]. Genetic factors seem to influence gene segment frequencies within the final naïve B cell repertoire, as V and D segment-use profiles are shared to a greater extent between twins than among unrelated individuals [14]. The CDRH3 sequence composition of B cell repertoires is however rather personal, even for individuals with identical genetic background [14]. Although they did not cover the whole B cell repertoire with their technique, the authors from this study conclude, that the system is able to generate a greater diversity than can be represented by the actual circulating B cells [14]. This would indicate, that even individuals with the same genetic background might react differently to common antigenic challenges [14]. Although V and D segment usage is biased, the joining of V, D and J gene segments [14] as well as heavy and light chain pairing [15,16] seem to follow a stochastic linkage according to their abundance in the final repertoire.

Not all the rearrangements are retained within the final naïve B cell repertoire. From deep sequencing data it could be inferred, that 69% of the initially rearranged IgH sequences are deleted from the repertoire of the final naïve B cells, with a selection against sequences containing long CDRH3s and hydrophobic patches [17]. These characteristics have been associated with autoantibodies but broadly neutralizing HIV antibodies containing such features have also been found [17]. This could be an explanation for why it is so difficult to consistently elicit such antibodies.

As for many other body functions, also the human B cell immune response is affected by ageing [18,19]. The frequency of autoantibodies increases with age [19]. Furthermore, a poor health status of elderly could be associated with a reduction in B cell repertoire diversity [20]. Since peripheral blood is composed of several different B cell subsets, a deeper analysis is required to completely understand this phenomenon [21,22].

1.1.2 Different B cell subsets with specialized functions

Several studies estimated the number of B cells that are circulating within the human body with varying numbers [23-25]. These slight discrepancies between studies are probably due to the biological differences between individuals. It is estimated that of the 10^{11} B cells in the human body [26], on average $2x10^9$ [24,25] with a median of $3.1x10^8$ [23] are circulating within the peripheral blood. Furthermore it has been estimated, that the latter B cell subpopulation is composed of $2\text{-}9x10^6$ different CDRH3 sequences [27]. The naïve B cells ($CD27^-IgM^+IgD^+$), which have never encountered an antigen, make up 60-70% of the peripheral blood B cells [22]. In the absence of antigenic stimulation they die after several days of circulation between blood and lymphoid tissues, so that they need to be continuously replenished from the bone marrow with a considerable turnover over time [22]. The remaining B cells that are patrolling our body can be subdivided into several experienced and/or specialized B cell subsets [28].

Based on findings from mouse studies, B cell antigens have initially been subdivided into T cell dependent (TD) and T cell independent (TI), with a corresponding subdivision of mature B cells into three major subsets [29,30]. Follicular B cells - which can also respond to TI antigens - are giving rise to long-term memory B cells and antibody secreting plasma cells in response to protein antigens that elicit strong TD germinal center responses. Marginal zone B cells and B1 cells on the other hand respond more rapidly to TI antigens and are thought to give a first line of defence until the more specific germinal center response is taking over. TI antibody responses have further been subdivided into TI type 1 and TI type 2, with the former being induced by B cells that have been activated through Toll-like receptor (TLR) signaling. The latter are induced by antigens containing highly repetitive elements, which activate B cells through extensive crosslinking of their B cell receptor. This shows, that the B cell immune response to pathogens is a complex process involving different B cell subsets

and differentiation pathways, which can be influenced by the nature, the concentration and the location of the antigen [30,31]. Both innate and adaptive immune cells can influence the outcome of B cell immune responses [30,32,33], but also the B cells themselves can take over regulatory functions [34,35]. For this reason, this rather simple classification of B cells into these different subsets has recently been challenged [30]. Since pathogens are composed of a mixture of antigens of different natures, the B cell response towards these immune stimuli is very likely composed of a mixture of different pathways.

Important to consider are also differences between the B cell immune systems of mice and humans, which means that findings in the mouse model always need validation before being considered also true for the human situation [36]. As opposed to mouse B cells, human B cells for example do not respond directly to lipopolysaccharides (LPS) [37]. All attempts to find the human counterpart to mouse B1 cells - which are thought to arise from a different origin than marginal zone and follicular B cells (collectively also referred to as B2 cells) and are mainly found in peritoneal and pleural cavities from which they secrete natural antibodies in the absence of antigen - failed so far [38-41]. There is evidence, that the most recently proposed human counterpart ($CD20^+CD27^+CD43^+CD70^-$) [38] actually represents B cells that are in a pre-plasmablast stage rather than B1 cells[41], however the debate about these cells is still ongoing [42]. Although CD27 was found to be a good marker for memory B cells in humans [43], the mouse field had more difficulties to find mouse markers that clearly distinguish mutated from unmutated B cells [44]. For splenic marginal zone B cells (sMZcs) – which play a major role in the response to blood-borne pathogens – the human counterpart is also still not clearly defined and controversially discussed. As opposed to the mouse, these cells harbour mutated V genes and recirculate between blood and the spleen [24,28,45]. Currently it is hypothesized, that the human sMZcs arising early in life are probably generated in an antigen-independent manner to form the pre-immune B cell repertoire, but that later in life, these cells predominantly arise from antigen specific germinal center responses [28,45].

An overview of B cell markers used in this study to differentiate B cell subsets can be found in Table 1.

Table 1: List of important B cell markers used in this study to analyze B cell subpopulations.

Marker	Description
CD19	Pan-B cell marker [22,46]
CD20	Pan-B cell marker not expressed on plasmablasts/plasma cells [22]
CD27	Memory B cell marker highly expressed on plasmablasts/plasma cells [46]
IgD	Together with CD27 allows to distinguish naïve B cells and non-switched memory B cells from classical memory B cells[46,47]; an association was made between $CD27^{-}IgD^{-}$ B cell numbers and systemic autoimmunity as well as chronic infections [47]
IgM/ IgG	Allow to determine isotype expressed on memory B cells [46]
CD38	Expressed at high levels on plasmablasts and plasma cells [46]
CD138	Marker for plasma cells [46]
CD23	Activation marker; involved in T-B cell interactions [22,48]
CD5	Initially thought to be the marker for the human counterparts of mouse B1 cells, it is now known that this marker is expressed on a wide number of different B cell populations [49]; CD5 is also expressed on malignant B cells [50]
CD10	Marker that is expressed on transitional B cells [51] and on germinal center (founder) B cells [52]
CD24	Combined with CD38 this marker allows to distinguish transitional B cells from mature naïve B cells [47]
CD43	This marker was shown to be expressed on the human counterpart to mouse B1 cells [38,49] but more recently this view was challenged and it was proposed that $CD43^{+}$ cells are rather B cells that entered a pre-plasmablast stage [41]
CD21	It was found that $CD21^{low}CD38^{low}$ B cells might play a role in autoimmune diseases and immunodeficiencies [47]

1.1.3 Diversification of B cell receptors during the B cell immune response

Proteins capable of activating both T helper cells and B cells elicit highly specific germinal center responses [53,54] which allow the generation of long term protective memory (Figure 2). While it is important to keep in mind that other pathways exist [30,55], due to their optimization for binding to the antigen, B cells generated by germinal center responses are most interesting for biomarker research. Inside these structures, selected B cells undergo several rounds of B cell receptor somatic hypermutation (10^3 per base pair per generation versus 10^6 under normal conditions with cell cycle times of 6-12 hours) and affinity selection to modify their receptor for a better recognition of the antigen[54]. These processes are responsible for the increase in serum antibody affinities that have been observed to occur over the course of immune responses [56-58]. Although B cell receptor mutations can be found all over the sequences, they seem to be targeted to hot-spot motifs (cold-spot motifs also exist) [59], an information that can be very useful for the interpretation of B cell receptor sequencing results [60]. Due to easier accessibility of lymphoid tissues and the possibility of genetic manipulations, a detailed understanding of germinal center responses could be obtained mainly through animal experiments.

Since the generation of highly specific antibodies is rather a slow process – requiring the establishment of germinal center structures in which several rounds of somatic hypermutation and selection occur -, at initial stages of the immune response a large fraction of activated B cells directly differentiates into short lived extrafollicular plasma cells [61-65]. Both a higher affinity of the B cell receptor for the antigen but also a higher epitope abundance preferentially induce B cells to undertake this initial short-term plasma cell pathway [64,65]. Within the first week of the response, these unmutated cells expand in the lymphoid tissues and give rise to a first line of defence until the more specific and long-lived plasma cells generated in the germinal center responses can take over [56,61,63,66]. From the large amount of initially formed plasma cells, only those with the highest affinity for the antigen survive as long-lived antibody producing cells in the bone marrow [56]. Interestingly in human studies it was observed, that besides antibody secreting plasmablasts specific for the immunizing antigen, a second antibody secreting plasma cell subpopulation of unknown specificity appeared in peripheral blood [67]. This indicates, that as a result of limited space inside the bone marrow, newly formed plasma cells might induce resident cells to leave their niche. Since long-lived plasma cells are the cells of the B cell immune response that are most specific for the antigen, their main role is to protect the body against the same pathogen through the action of their secreted antibodies [68].

A second product of germinal center responses are long-lived memory B cells, which allow our body to elicit much faster and more efficient B cell immune responses upon reencounter of a similar immune stimulus [28]. These cells seem to be less affine for the antigen [56,69], meaning that they probably have a greater polyspecificity allowing also protection against related or mutated variants of the same pathogen [68,70]. From a biological point of view, it makes sense to first generate the antibodies that confer protection against the pathogen that invaded the body and only later produce memory B cells that can also confer protection upon reinfection with a similar pathogen [56,70]. This does however not exclude the possibility, that some early memory B cells could be formed outside of germinal center responses [71]. Generally the many discrepant results in the field regarding clear separation of B cells into subsets and determining their fate indicates that there is some degree of stochasticity in the system. Since so many factors are involved, certain B cell types might have a higher intrinsic preference to undergo a certain pathway, but this does not necessarily mean that they cannot or will never also use another one. Since not being perfect is a prerequisite for the evolution and survival of biological systems, it only makes sense to consider this assumption.

In vitro culture experiments using human B cells revealed, that IgM memory B cells more likely re-eneter the germinal center reaction, while IgG memory B cells have a tendency to rather differentiate directly into antibody secreting plasma cells [72]. Although it is still under debate whether this is true for all human IgM expressing memory B cells, current data favour the view, that IgM memory B cells are the early products of germinal center responses, while isotype switched memory B cells are rather generated at later timepoints [28]. This would be in line with the uniform switching that was observed in early mouse experiments [61] and the general switching pattern from IgM to IgG isotype of antibodies appearing over the course of immune responses [58].

Although it is accepted, that human memory B cells only expressing IgM or IgG isotypes are generated in germinal center responses, the origin of those IgM memory B cells that did not downregulate IgD is still under debate [28]. Current data indicate however, that at least a large fraction of the $CD27^{+}IgM^{+}IgD^{+}$ memory B cells are the products of early T cell dependent responses [21,24,60,72-78]. A germinal center independent origin of these cells was proposed based on the finding, that these cells express B cell receptors and harbour mutations that are different from germinal center derived B cell subsets present at the same timepoint [21,24,73,74]. Since B cell receptors are largely changed in the course of germinal center responses and certain amino acids are

preferentially used to mediate contacts with antigens [79-84], a germinal center independent pathway needs to be confirmed by more direct approaches. Analysis of the evolution of these clones between different timepoints and sequencing of repertoires at greater depth [60,77] are required to completely solve this issue. It should also be considered here, that maybe not all the B cells entering the germinal centers at initial stages of the response will survive and enter the B cell subpopulations generated later during the response as mutated and class-switched versions. Transcriptome analysis combined with *in vitro* functional assays [72] and epigenetic data [78] indicate that $CD27^{+}IgM^{+}IgD^{+}$ B cells share more features with real memory B cells than with naïve B cells and are in line with the idea of these cells being generated early in T cell dependent immune responses. Furthermore, the increase in mutation frequencies from $IgM^{+}IgD^{+}$ to $IgM^{+}IgD^{-}$ and $IgG^{+}IgD^{-}$ B cells [24,74,77,78] is in line with the former cells being generated before the other subpopulations.

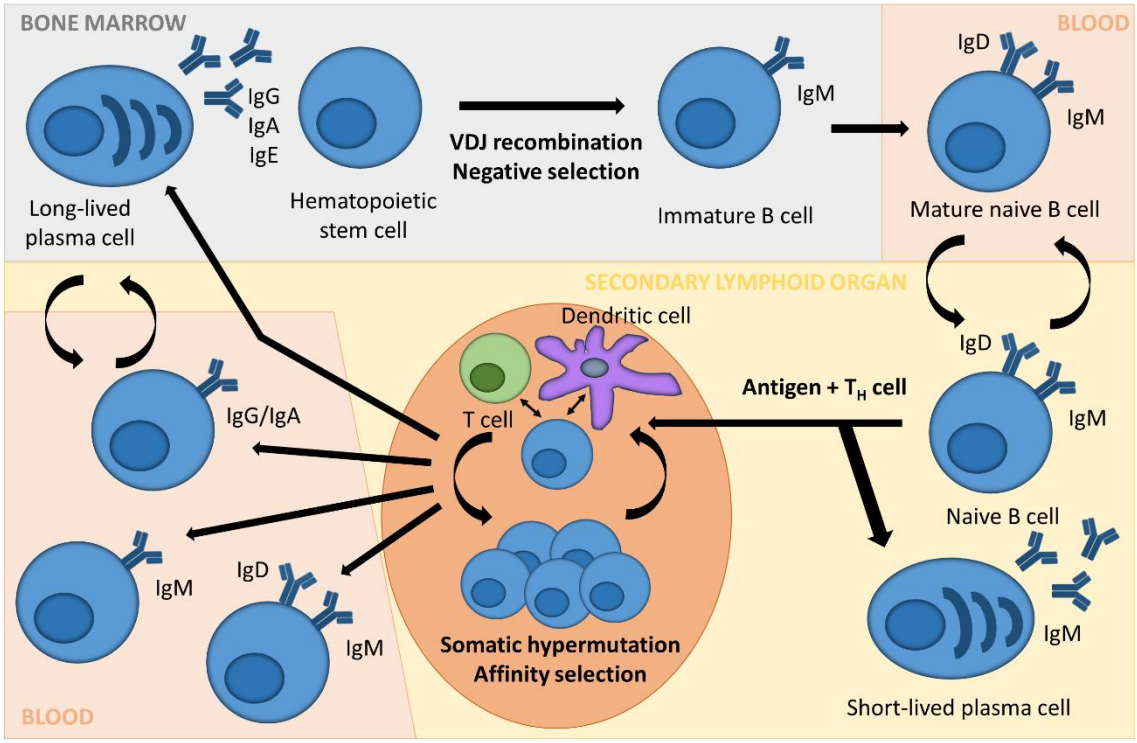


Figure 2: Diversification of B cell receptors during the immune response occurs in germinal centers. *Several rounds of somatic hypermutation and affinity selection result in the production of cells that have a high affinity for the antigen. In a first step B cells have to compete for access to signals from T helper cells before they vie with each other to get access to antigen presented by dendritic cells. Depending on the signals they get from these two cell types, B cells either continue proliferation, somatic hypermutation and affinity maturation in*

germinal centers or they differentiate either into long-lived plasma cells or memory B cells. B cells which do not get any survival signals will die. Only B cells with the highest affinity for the antigen will be selected to enter the bone marrow as long-lived cells that secrete antibodies to protect the organism against the pathogen that elicited the immune response. Memory B cells are generally less specific for the antigen and thereby allow to mount strong and fast immune responses also to related pathogens. Since the germinal center response is rather a slow process, a large part of B cells directly differentiates into short-lived plasma cells to provide a first line of defense until the more specific cells generated in germinal centers can take over.

In summary, T cell dependent B cell responses induce germinal centers in which B cells expressing receptors recognizing the pathogen are continuously selected to be optimized for binding and to enter the long-lived memory B cell compartments. These structures allow to further diversify the B cell repertoire in order to be armed for a potential rechallenge with the same or a similar pathogen. The idea that besides their role as inducers of highly specific and long-term antibody responses, germinal centers are also enlarging the B cell repertoire diversity in order to protect against mutated versions of the same pathogen [68,70], is an important consideration when analyzing B cell repertoire data after immunization or infection with the aim to identify specific B cell repertoire signatures. It appears that over the course of the immune response, germinal centers precisely switch their output [85] to send the cells of different affinities to the different B cell compartments fulfilling precise functions. Similar to the effect seen when combining different drugs to target the same pathogen, the generation of a diverse antigen-reactive repertoire [70] might also prevent the occurrence of escape mutants. It is interesting to mention in this context, that more diverse B cell repertoire responses could be associated with better survival after infection with the influenza A (H7N9) virus [86].

1.1.4 B cell receptor selection at different stages of the B cell immune response

Competition between B cells possessing different antigen affinities is dictating the different stages of antibody affinity maturation and selection [64,65,87,88]. This means, that whether a B cell with a particular B cell receptor is responding or not can depend on the presence or absence (both with regard to time and location) of other B cells expressing B cell receptors with a different specificity for the antigen. In the absence of B cells with high affinity, the less affine ones will take over, while in a system where more affine ones are present, the same B cell might not respond.

Germinal centers have been shown to be open and dynamic structures [89,90], meaning that new B cells can constantly enter and influence the outcome of pre-established germinal center responses. This is very important, when considering complex antigens [91,92] or pathogens like *Borrelia burgdorferi* that can even change antigen expression in the course of immune responses. This means, that not only the presence or absence of a certain B cell receptor but also of epitopes can influence the outcome of the B cell response at the B cell receptor level. This also means, that initial responses are probably dominated by B cell receptor sequences that are highly abundant in the pre-immune repertoire but that these might later be exchanged by other more rarely occurring ones allowing higher affinity binding [93,94].

Mouse studies have revealed, that although the naïve B cell repertoire composition is largely influenced by genetic factors, the selection of B cell clones into the plasma cell pool is rather occurring stochastically [95]. This rather stochastic selection does however not seem to prevent a higher clonal overlap among plasma cells from individuals that have been challenged with the same immune stimulus [95].

To avoid uncontrollable spreading of pathogens within the body, the speed at which immune responses happen is crucial and explains the need for the above described efficiency of the system. In order to allow for a quick recognition of the antigen, the naïve B cell repertoire is composed of highly polyspecific B cell receptors (Figure 3). Experiments using phage display libraries, were able to show, that individual germline antibodies can not only recognize a very large number of different epitopes [96], but that also a very large number of different germline antibodies (>1000) can recognize the same antigen [97]. The recruitment of B and T cells to special anatomical sites like the lymph nodes increases their chances to meet, thereby also speeding up the process [98]. Previous generation of an expanded and diverse memory B cell pool largely located to these sites might also contribute to a quicker response. Besides these strategies, it was proposed that B cells and T cells must undergo already some degree of proliferation before cognate interactions so that they can find their interaction partner more easily [98]. The degree of polyspecificity of generated memory B cells needs (to my knowledge) still be determined. In mice it could however be shown, that infection can induce the production of memory B cells that are capable of recognizing mutant variants even better than the viral strain that was used for immunization [68]. The fact, that B cell immune responses generally induce some antibodies that can also bind to unrelated antigens [99,100], means that

polyreactive memory B cells are also very likely generated. For this reason, the possibility that some memory B cells might participate in the immune response needs to be considered especially when analyzing human B cell repertoire perturbations.

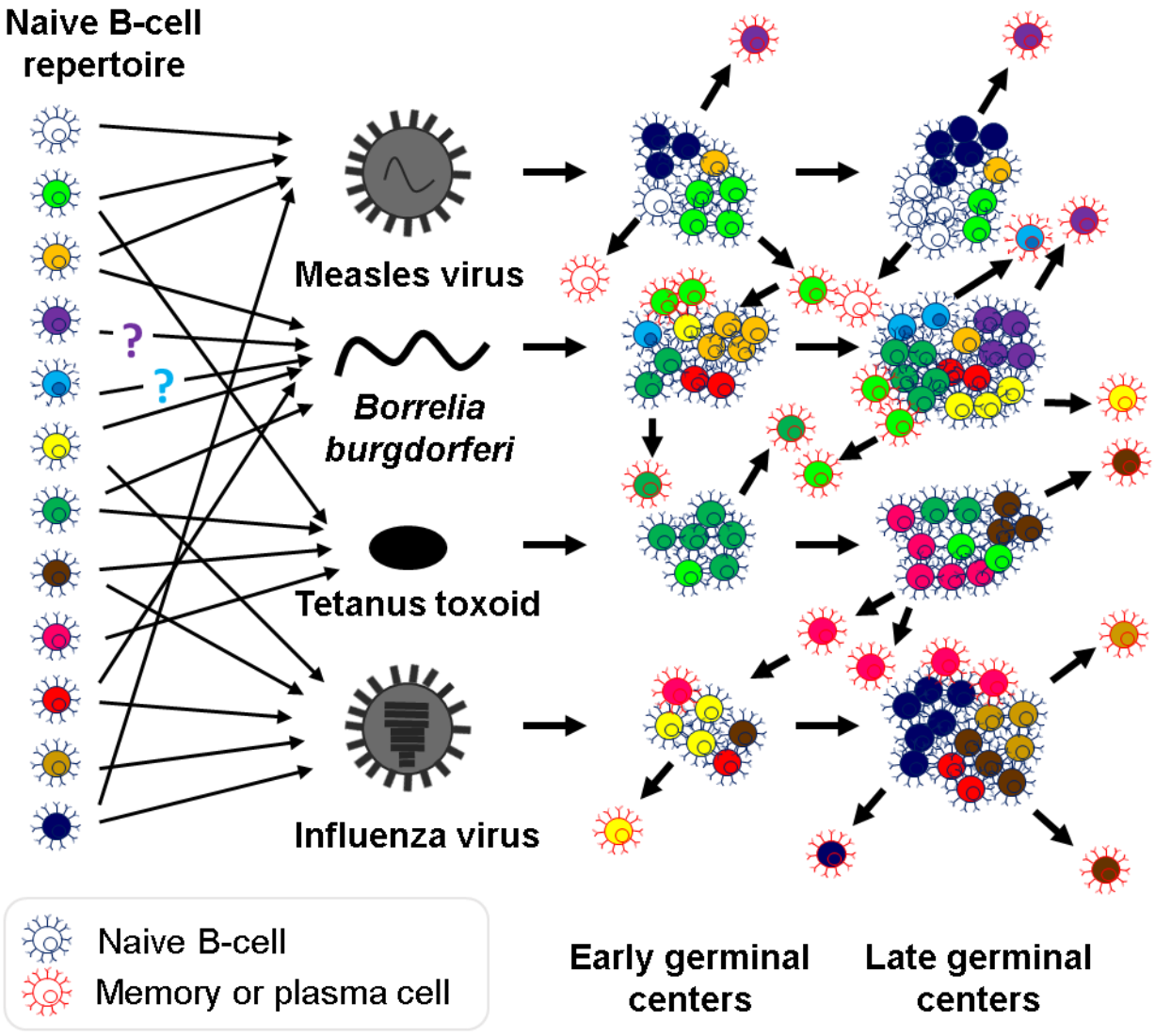


Figure 3: Selection of B cells into and out of the B cell immune response. *In order to allow for a quick B cell response, the naïve B cell repertoire is composed of B cells harboring polyreactive B cell receptors. These seem to be selected rather stochastically to enter the B cell immune response. As germinal centers are open and dynamic structures, new B cells can constantly enter and change the response and might lead to a switch in germinal center outcomes. Even if B cells are selected rather stochastically into the B cell immune response, some degree of convergence seems to occur. This means that similar B cell receptors might be generated in different individuals. Especially in humans it is important to consider, that memory B cells generated in previous immune responses might participate in the B cell immune response of interest. The main question of this project: can we identify among the diverse repertoire that is expected to react with Borrelia a few B cell clones that are specific for that pathogen and which are generated in common between different individuals (question marks).*

Due to the isolation and use of monoclonal antibodies in biological assays that allow to even distinguish between proteins that only differ by one amino acid position (for example: Thy1.1/Thy1.2 system [101]), biologists tend to have the reductionist view [102,103] of the antibody-antigen recognition process as being highly specific. In addition to that, the separation of protein sequences into defined antigenic regions [104] may lead to the wrong assumption, that different individuals recognize the same antigen in a similar manner. Even if a similar antigenic region is targeted, this does not necessarily mean, that this region is bound by the same mechanism [102]. Early studies already concluded, that immune stimuli induce the production of several different antibodies, which together render sera specific for that antigen but some of which might also be responsible for the frequently observed increase in cross-reactivity of sera with unrelated antigens [99,100]. Since only B cells harbouring the most specific antibodies are kept in the bone marrow to secrete antibodies, the timepoint of sampling might influence ELISA results. While one might find a large number of different antibodies reacting with the same antigen early in the response, later response sera might be predominated by only a few highly specific antibodies. Detailed investigation of the composition of human serum antibody repertoires by combining mass spectrometry with next generation sequencing is a new field that will allow to address these questions more thoroughly [105]. A first study using this approach was able to follow changes in serum antibody composition at different timepoints after tetanus vaccination [106]. Although a much larger plasmablast clone number was expanded in day 7 PBMC samples, only a few of the corresponding antibodies could be found back in sera of these donors at later stages of the response [106]. For the two donors analyzed in this study, only three clonotypes were responsible for over 40% of the measured antibody response [106]. Taken together with the rather large number of memory B cells being generated, this means, that at the cellular level the B cell immune response is way more diverse than at the antibody level (the latter can be in the range of 100 clonotypes [106,107]).

Antibodies generated in prior immune responses can also affect germinal center outcomes [108,109]. In the light zones of germinal centers, B cells first compete for uptake of antigen presented as antibody-antigen complexes by follicular dendritic cells and after that for helper signals from T cells which sense the amount of antigen presented. Antibodies that are produced in the course of the immune response shield epitopes on the antigens that are presented by follicular dendritic cells, thereby making sure, that only B cells with B cell receptors of higher affinity than that of already existing antibodies can get access to the antigen and thereby also to the required survival signals. This

is a way of how germinal centers can communicate with each other by other means than via cell migration and explains why B cell responses can stay so diverse without polarizing to a specific antibody over time. It seems though, that antibody feedback is mainly mediated by antibodies of the IgM isotype.

Although not all these studies directly selected for B cell clones reactive with the same antigen, sequencing of induced memory B cells or plasma cells from patients with the same disease conditions generally leads to the isolation of a rather broad repertoire composed of various different V-J gene combinations [110-116]. These studies are in line with a rather broad spectrum of B cell clones generated upon immune stimulation and also explain why many B cell repertoire studies showed, that overlapping clones between patients are very rare events [1,13,86,117-120]. Due to this great diversity, next generation sequencing techniques [121] are indispensable for the study of the B cell immune response at the B cell receptor level. Although this field matured extensively over the past years, even this technique only allows to analyze a small part of an individual's whole B cell repertoire [122]. For this reason, future studies will need to combine next generation sequencing with single cell isolation approaches [123,124] to adequately characterize and determine the sizes of antigen-reactive B cell repertoires.

1.1.5 The kinetics of B cell immune responses

Although at the B cell receptor level, the response seems to be largely influenced by stochastic and environmental factors, general kinetics of the B cell immune response are surprisingly similar between different immune stimuli [4] suggesting the existence of some buffering mechanisms [125]. Generally it can be assumed, that around one week after vaccination or onset of symptoms, large numbers of plasma cells are released into the circulation. They are present there at elevated numbers however only during one day. It is interesting to mention in this context, that the peak of the extracellular plasma cells responsible for initial antibody production is occurring at exact this timepoint in mice [61]. Although these experiments showed, that a large fraction of these B cells are eliminated by undergoing apoptosis [63], this coincidence in time raises the question of whether this striking rise in peripheral plasma cells is the result of a rather polyspecific early first line plasma cell burst. Many human studies have focused on this early plama cell peak for the isolation of antigen-reactive plasma cells. It seems though, that only a small portion of those is contributing to the long-

lived serum antibody repertoire [106]. In humans, plasma cells at this stage already harbour a high number of mutations, which is why a large fraction is probably resulting from pre-existing memory B cells that can have been generated in previous responses to related or unrelated antigenic challenges [116,126-129].

For long-term survival, antibody secreting plasma cells need to find a niche inside the bone marrow, from which they continue secreting antibodies that can be detected for prolonged times in the serum [130-134]. Upon immune stimulation, memory B cells also increase, although at a much slower and lower levels than the plasma cells [116,135-138]. The fact that for over one month period and only after immunization, PBMCs can be stimulated with the corresponding antigen to secrete antigen reactive antibodies [139-141], indicates that antigen activated and reactive B cells are circulating in peripheral blood over a longer period. Those cells are of particular interest for biomarker research. Furthermore older studies were able to show a slight increase in peripheral blood B cells upon acute infection [142,143], further indicating that they could be of use to support diagnosis [4]. As opposed to antibody secreting plasma cells – which home to the bone marrow -, memory B cells largely localize to secondary lymphoid tissues and although they largely decrease in numbers after clearance of the antigen, some continue circulating between these locations and peripheral blood for a lifetime [23,133,144,145]. The continuous evolution and increase in antibody affinities over time [56-58] suggests that also some plasma cells must be continuously generated. A low level of plasma cells is also present in peripheral blood in the absence of infection or vaccination. These cells were shown to be derived from constant immune responses going on at mucosal sites [146].

It is assumed, that a decrease in antigen towards the end of immune responses is leading to an increase in competition between germinal center B cells thereby allowing for the gradual selection of more and more specific clones [147]. The response is believed to stop, when the antigen has dropped below a certain concentration [125]. An interesting and important question to raise here is: what happens in the case of chronic antigenic stimulation? The finding, that a substantial number of reactive B cells can localize directly inside chronically inflamed tissues [117,148] raised the question of whether reactive B cells can be reliably isolated from peripheral blood samples. Examples describing elevated numbers in the periphery however also exist. In the autoimmune disease systemic lupus erythematosus – in which the autoantigen is constantly present -, a correlation between disease activity and peripheral blood plasma cell levels could be observed [149,150]. Also the analysis of

peripheral blood samples from a chronic HIV infected individual allowed to follow the co-evolution of viruses and anti-HIV broadly neutralizing antibodies over a three years time period [151]. In multiple sclerosis shared clones between the central nervous system and peripheral blood could also be found [152]. These findings indicate, that antigen-reactive B cells should not only be found in peripheral blood upon acute but also upon chronic infection making them attractive candidates for biomarkers to indirectly show the presence of pathogens inside the body. It seems though, that the plasma cells generated under chronic auto-immune stimulation are of a different nature than the long-lived plasma cells generally generated upon single booster immunizations [153,154]. In mouse models, it was found that chronic antigen stimulation and inflammation are inducing the generation of short-lived plasma cells [154]. A very interesting and important finding from this study was, that chronic inflammation seems to induce long-lived bone marrow plasma cell displacement and prevents entry of newly generated plasma cells into the long-term survival niches of the bone marrow [154]. If it can be confirmed, this means that chronic inflammation might affect protection against previously encountered immune stimuli. It could also be shown, that injection of soluble antigen into an ongoing germinal center response leads to apoptosis of antigen-reactive germinal center B cells, a strategy that might be exploited by some pathogens [155-157].

1.1.6 B cell repertoire signatures

As already mentioned above, the great diversity and polyspecificity of the naïve B cell repertoire in combination with the finding, that antigen reactive B cells are often composed of various different VJ gene combinations, suggests that B cells are rather stochastically selected from the primary repertoire to start the immune response. It seems though, that some degree of convergence is occurring especially at later timepoints during which germinal centers should have generated B cell receptors of high specificity. This suggests, that there are probably some optimal B cell receptors in the repertoire for each antigen, but since the body needs to respond very quickly, as a first line of defence at early timepoints, he is working with the best B cells available at that time. Later in the response more B cells have circulated the affected lymphoid tissues and - since germinal centers are open structures - might lead to the appearance of more similar sequences among individuals. In addition to that, their B cell receptor is optimized for binding in germinal centers, and the selection of similar amino acid changes might also lead to a higher convergence. This is what one expects to happen in a primary immune response, but the situation might be different when memory B cells are also involved.

Haptens coupled to carrier proteins and hen egg lysozyme are two groups of antigens that were extensively studied in mouse models. Very early studies already addressed the question: how diverse is the antigen-reactive B cell repertoire, and their results are in line with more recent findings. In the 1980ies a common signature characterizing B cell receptors of high affinity for the hapten NP was identified [158]. It was found, that one amino acid change in the V-gene (tryptophan to leucine exchange at codon 33 in the CDR1 region of the mouse V186.2 germline sequence) was repeatedly occurring and responsible for a 10 fold increase in antibody affinity for the antigen [158]. The identification of this signature allowed to analyze and follow the fate of germinal center B cells in great detail [56]. It seems though, that these clones are not the only ones generated in response to this vaccination. Especially secondary immune responses seem to induce a broader antigen-reactive repertoire composed of different V genes [94,159]. Analyzing the *in vivo* adaptation of the anti-henn egg lysosyme monoclonal antibody HyHEL10 to a mutated version of this antigen (HEL^{3x}) also indicated that a key mutation (substitution of tyrosine at codon position 53 with aspartate) was responsible for the observed increase in antibody affinity (~85 fold) [66]. This mutation occurred in over 95% of the isolated high affinity clones [66].

Due to their great potential for the generation of universal vaccines, phylogenetic pathways leading to the generation of broadly neutralizing antibodies to HIV and influenza have been studied extensively and contributed significantly to a better understanding of antibody-antigen interactions and their co-evolution. Even if it is important to keep in mind that these are rather special cases, several interesting conclusions regarding B cell repertoire signatures can be drawn from these studies. Approximately 20% of individuals generate broadly-neutralizing antibodies upon HIV infection. In one fifth of these patients, antibodies of the VRCO1-class could be identified [160]. These antibodies are generated from a common germline V-gene (IGHV1-2*02) and harbour an exceptionally high number of mutations [118]. Even though these sequences only share less than 50% identity, they seem to bind to their target through similar mechanisms [160]. Also in this case, key mutations at two residues in the CDR2 region are the only globally shared features of isolated heavy chain sequences (glycine to alanine at position 56 and threonine to valine at position 57) [118]. Following the ontogeny of broadly influenza neutralizing antibodies showed how tremendously single amino acid substitutions can change binding to antibody targets [161]. This suggests, that antibodies with very different sequences can adopt very similar three dimensional antigen binding mechanisms and that key residues rather than whole sequences might be the signatures we are looking for in B cell

repertoire data. This is also in line with more recent computational analysis, that lead to the conclusion, that only a few key residues are responsible for the majority of the difference in energy that occurs upon binding to the antigen [84]. It is important to emphasize in this context, that two different types of amino acid changes occur in antibody sequences. One part of these amino acids might directly interact with the antigen, while the other part only indirectly contributes to binding by affecting the overall antibody structure [6].

The fact, that only a subset of patients is capable of forming broadly neutralizing antibodies indicates, that the B cell immune response at the B cell receptor level can be highly personal. Broadly neutralizing antibodies have been shown to have a higher degree of autoreactivity, which is probably the reason why they are occurring rather sporadically [162]. As mentioned above, the convergent nature of these types of antibodies is thought to be rather an exceptional case [119]. Since viruses tend to mimic host proteins to avoid recognition by the immune system, it has been hypothesized, that the hydrophobic IGHV1-69 gene – which predominates influenza broadly neutralizing antibodies - has occurred during evolution because it allows to bind to important hydrophobic regions on viruses (including influenza, HIV and HCV) [119]. Due to their higher risk of generating auto-immune reactions, B cell receptors containing highly hydrophobic HCDR3s are however generally deleted from the repertoire [17]. This means, that usage of this particular V-gene might be the only possibility to generate such antibodies, which is why if they occur, they are highly similar in different individuals.

In accordance with the above statements, pathogenicity in pemphigus patients was associated with the presence of one tryptophan in the CDR3 of the heavy chain of isolated Dsg-specific antibodies [163]. Another example for which a disease signature was identified by looking at the B cell repertoire is multiple sclerosis. In that case, an enrichment of VH4 expressing B cells harbouring replacement mutations at eight defined codons was found in cerebrospinal fluid [164].

Even if some signatures could be identified, the great diversity and polyspecificity of antigen reactive B cell repertoires suggests, that the B cell immune response is a highly personal process largely governed by stochastic and environmental factors. For this reason, the isolation of disease specific signatures that would allow to discriminate between different conditions was initially thought to be rather difficult. The CDR3 sequence of the heavy chain was shown to be most implicated in antigen

binding and is sufficient to confer specificity of an antibody for its antigen [11]. For this reason, one would expect convergent signatures to occur rather at this site. The only condition however for which a clear convergent CDR3 sequence could be associated with disease activity is dengue [165]. Since specificity of antibodies with this CDR3 for dengue was never proven experimentally, the question occurs of whether this high degree of convergence might be explained by the polyreactive natural B cell activation that was shown to occur in this disease [166]. On the other hand, the generation of a highly diverse antigen reactive repertoire does not exclude the possibility that some sequences nevertheless might be generated in common in different individuals. It might just be a matter of how deep one digs into the repertoire, a problem that can be overcome by using next generation sequencing to analyze B cell immune responses [167,168]. With this technique, several groups were able to show a higher convergence of repertoires among individuals exposed to the same antigenic challenge [127,169-171], meaning that overlapping clones should exist. If the occurrence of overlapping clones among different individuals suffering from the same disease condition can be confirmed, this would make B cells very attractive biomarkers for diagnosis. It remains to be investigated though, how specific these overlapping sequences are for each disease condition and how globally they are occurring among individuals. Other important considerations are their kinetics and levels in peripheral blood. Recent studies indicate, that the overlapping clones might predominantly occur at low levels during later stages of the immune response [171] raising doubts about their usefulness for diagnosis. Since every antibody-antigen interaction as well as every B cell immune response to a certain disease is rather unique, it might still be worth investigating each disease individually. Our current knowledge of the B cell repertoire is not big enough to draw global conclusions. It could very well be that for certain diseases, specific signatures can be isolated, while for other conditions this might not be possible.

Even if B cell repertoire deep sequencing studies might not lead to the isolation of convergent signatures, these data contain a tremendous amount of other interesting information, which allows to better understand the B cell immune response to various conditions. Conclusions from such studies can be very valuable for the development of vaccines or to understand why certain patients are protected while others are not and why some are more prone to develop chronic courses or autoimmune diseases.

1.2 Lyme disease and it’s causing agent

1.2.1 Lyme disease and problems encountered with diagnosis

With an estimated 65500 cases per year in Europe [172] and 300000 in the United States [173], Lyme disease is the most common tick transmitted disease in these two regions of the world. One needs to keep in mind though, that these numbers are just estimates and that real numbers might largely deviate from these estimations. The disease is caused by a bacteria, or more precisely a spirochete, that resides inside the midgut of ticks [174]. The latter are blood feeding ectoparasites that belong to the same arthropod class than spiders and mites (Arachnida) and that - similar to mosquitoes - can transmit various different diseases during blood feeding [175,176]. Although other tick species exist, the most important vector for Europe is *Ixodes ricinus* [174]. As opposed to mosquitoes, the blood meal of this tick generally takes longer. Ticks developed different mechanisms to overcome the host immune response which would otherwise attack them and prevent them from feeding [177-180].

Borrelia generally first need to migrate from the midgut of the ticks to their salivary glands, from which they are then secreted into the skin of the host [181-183]. For this reason, it is assumed, that they are usually only transmitted after 24-48 hours [184]. Differences between different *Borrelia* species were however observed, with some species being transmitted earlier [184,185]. Since no vaccine is currently available [186], avoidance of areas with high tick incidence and early correct removal of these parasites is the best prevention strategy of the disease for humans [174]. Although it was shown that disturbance of the tick – for example by adding nail polisher or by squeezing it - did not increase infection rate [185], it is still recommended to not harm the tick upon removal. Prophylactic antibiotics treatment [187-189] might be another option to prevent *Borrelia* dissemination inside the body, but this is not an optimal solution especially in the case of frequent tick bites.

Due to their transmission route, *Borrelia* enter the body of the host through the skin. The helical shape of this bacterium allows it to migrate at a very high speed inside tissues [183,190]. After their deposition, the bacteria migrate away from the site of the tick bite and can invade the body. Motility of the bacteria is important to establish an infection [183]. In 60-80% of patients, migration through the skin leads to the characteristic early manifestation called erythema migrans [191]. This symptom is characterised by the expansion of a red circle around the tick bite site and can be explained by the

immune response of the host following the migrating bacteria [190]. Not every patient develops this characteristic symptom [192-195] and also tick bites are frequently unnoticed [196]. Misdiagnosis of this early manifestation has also been observed [197]. In case the patient does not manage to combat the bacteria at early stages, they might disseminate to different parts of the body including other sites of the skin, the nervous system, the joints or even the heart [174,193,198-200]. A study performed in Northern Europe was able to show, that within three months after the tick bite 40% of infected individuals develop disease manifestations [192]. Since *Borrelia* are not known to produce toxins, the symptoms are probably a result of inflammation caused by the infection [174]. Lyme disease has many different faces [174,193,198-200] and many symptoms are similar to those observed in other diseases [201]. This can render diagnosis difficult, especially when no erythema migrans or tick bite was noticed. Even though the diagnosis might be easy for specialists in the field, it might be complicated for general practitioners which only rarely see Lyme disease cases [202]. For this reason the development of diagnostic tests allowing a clear yes or no result is highly important. Current diagnostic procedures predominantly rely on the correct identification of patient symptoms and exclusion of other possible diseases. In the absence of erythema migrans, serological tests should be used to confirm the diagnosis [173,203-207]. One needs to keep in mind though, that a positive serological result can only indicate whether the patient has once had contact with *Borrelia* but cannot give any indication whether the infection is still acute [173,208]. IgM antibody testing is only recommended to support diagnosis of early stage disease (first month after tick bite) in the absence of a clear erythema migrans [173,204,205,207]. Current improved versions of serological tests should detect IgG antibodies in the majority of later stage patients [203,205] and in their absence, other possible reasons for the symptoms should be considered. Although serological tests have been extensively optimized in the past [203,207], scientific studies often only focus on one specific aspect of a problem and therefore do not allow to reproduce the real situation occurring in practice. In current recommended tests, an ELISA followed by a more specific immunoblot are used to confirm seroconversion. Problems with correct interpretation of immunoblot results in practice have been observed [207,209,210], which is why a two ELISA test system with an automatic readout has been proposed to solve these practical issues [207]. Issues with different performances of different tests are also not solved yet [209,211,212], so that a universal test system still needs to be established.

Since Lyme disease is caused by a bacteria, it can be treated with antibiotics and the majority of patients recover well from the disease [174,194]. In the majority of cases, the disease would even resolve in the absence of treatment, but it is recommended to give antibiotics to prevent a possible

dissemination of the bacteria that could cause later stage complications [174,193]. There is a subset of patients that continues to have symptoms after treatment, a phenomenon that is more frequently observed in donors suffering from later stages of the disease in which the bacteria had time to largely disseminate inside the body [174,213-215]. Several explanations exist. On one hand, prolonged infection and tissue inflammation might have caused some irreversible damage inside the affected tissues and might prevent the patients from full recovery. It was also suggested, that persistent presumably antibiotics resistant *Borrelia* might be the reason for these continuous symptoms. Clinical studies from the United States indicate, that additional antibiotics treatment might alleviate chronic fatigue in some patients, but patient improvement might be associated with a high risk of side effects, which is why prolonged antibiotics treatment is currently not recommended [216]. Although it cannot be completely ruled out, that in individual cases antibiotics treatment might have failed to completely eradicate *Borrelia,* especially in cases where a wrong treatment schedule might have been followed, currently there exists no solid proof for the occurrence of resistant *Borrelia* in humans [193,214,217]. Real life situations might also be different from the well selected patient cohorts used for scientific studies. It was for example found, that patients with haematological malignancies might require more often retreatment than immunocompetent individuals but also in these cases the outcome was good after treatment [218]. Reoccurrence of Lyme disease manifestations is generally caused by a different *Borrelia* strain, indicating that reappearing symptoms are due to reinfection rather than a relapse of the disease [219-221]. Since *Borrelia* tend to only be present at very low numbers in accessible human samples, it is very difficult to prove or disprove the persistence of these bacteria inside the patients. For these reasons *Borrelia* culture and PCR approaches are only being used for research purposes but are not reliable enough to be used as diagnostic tests of acute infection [173,202,203,222]. Even if these methods have a low sensitivity, one would still expect a positive result in individual antibiotics resistant cases, which has not been observed [223]. Of course it is hard to prove that there are no *Borrelia* anymore inside less accessible tissues. Since the bacteria seem to be able to sense the presence of a feeding tick, lately xenodiagnosis was proposed to test for the presence of viable bacteria inside chronically infected humans [224,225]. In this approach, germ-free ticks are put onto patients to feed and these ticks are later screened for the presence of bacteria. Although individual ticks from one post-treatment Lyme disease patient were positive by PCR for *Borrelia* DNA, they failed to isolate viable spirochetes [224,226]. Some studies showed the presence of cystic structures that react with *Borrelia* antibodies in tissues from patients, but no clear correlation between the presence of these structures and persistent symptoms after treatment has been established[227]. In addition to that, one antibody used in these studies seems to cross-react with human proteins, meaning

that confirmatory experiments other than immunological staining are required to confirm these findings [227]. A mouse study was able to demonstrate the persistence of *Borrelia* antigens after treatment, indicating that the body might have problems to eliminate all remnants of dead *Borrelia* and that this could potentially be the cause of the observed persistent inflammation [228]. Similar to this, the "amber theory" of chronic Lyme arthritis hypothesizes, that inclusion of *Borrelia* remnants into joint matrix could result in recurrent inflammation in case they are release into the joint space [229]. There is also evidence, that ongoing inflammation and autoimmune responses might be involved at least in a subset of patients. In some cases, autoantibody producing B cells could be isolated from affected tissues and the presence of autoantibodies in sera from patients have also been observed [230-247].The type of T cell help that B cells get might play a role in the chronic course of the disease. Patients developing higher levels of CXCL9 and CXCL10 chemokines are more prone to produce anti-*Borrelia* antibodies [248] and *Borrelia* are more difficult to isolate from these patients [223]. These Th1 cell responses seem to be important for efficient killing of spirochetes. An association was made between Th17 cell driven responses and the formation of autoantibodies [223]. This effect might not be directly visible, when simply comparing autoantibody levels between resolved and chronically evolved patients [249]. Over the course of the response to *Borrelia*, there seems to be however a shift in the Th17 immune response from a protective role early in the disease towards a detrimental one if not switched off at later stages (at least in the case of Lyme arthritis) [250]. It seems that a large part of patients develop transient autoantibodies, but only in a subset of these, these antibody titers remain high for prolonged times [237,249]. It could thus be, that depending on the signals that the B cells get, they might develop clinically important autoantibodies later in the disease [250]. The exact role these antibodies play in the different forms of post-treatment Lyme disease conditions remains to be more clearly investigated though[250]. General problems to clearly differentiate post-Lyme disease patients from resolved ones [249,251] indicate, that multiple and highly complex mechanisms might be responsible. Personalized medicine and systems biology approaches on very large patient cohorts will probably be required to solve these issues.

The chronic manifestations observed after Lyme disease infection are actually not specific for Lyme disease, but similar symptoms have also been observed under many other conditions [252-256]. These manifestations are frequent among the general population [193,210,257,258]. Due to the increased awareness and popularity of Lyme disease, it is often used as explanation in case such symptoms appear. This very likely leads to frequent misdiagnosis of chronic Lyme disease [213,259]. Only patients for which a previous acute Lyme disease infection can be proven should be diagnosed of

having chronic Lyme disease [213]. A less confusing and more correct term to use in this case would be "Post-treatment Lyme disease syndrome" [214] or "Post-Lyme disease symptoms" [213]. In the absence of an indication of previous acute Lyme disease, other reasons should strongly be considered.

1.2.2 *Borrelia* and important antigens they express

The tick-transmitted spirochetes of the genus *Borrelia* can be divided into two groups. One is transmitted by soft ticks (Argasidae) and is causing tick-borne relapsing fever, while the other one is transmitted by hard ticks (Ixodidae) and causes Lyme disease [260,261]. Louse-borne relapsing fever also exists. Tick-borne relapsing fever occurs only rarely in Europe but needs to be considered under special cases, especially when the patient has travelled to endemic areas [262]. Recently it was however found, that *Borrelia miyamotoi*, a relapsing fever spirochete present in the most important tick species complex from Europe [260], can also cause relapsing fever [260,263]. Since this spirochete can be transmitted transovarially, its importance for public health needs to be strongly considered [260]. Furthermore, antibodies induced by this group of *Borrelia* can cross-react with Lyme disease antigens, and might for this reason occasionally be important to consider for the correct interpretation of Lyme disease serological results [260].

Lyme disease causing *Borrelia* belong to the *Borrelia burgdorferi* sensu lato complex, which comprises over 20 different genospecies [264]. For simplicity, when I am referring to *Borrelia* here without any other indications, I mean this group of spirochetes. Not all the members of this group are human pathogenic. Pathogenicity could be confirmed for five species (*Borrelia burgdorferi* sensu stricto, *Borrelia garinii*, *Borrelia afzelii*, *Borrelia bavariensis* and *Borrelia spielmanii*) [193,264]. While in the United States, *Borrelia burgdorferi* sensu stricto is mainly responsible for human Lyme disease, the situation is more complicated in Europe. Here several different species can cause disease, *Borrelia afzelii* and *Borrelia garinii* being the most frequent ones [174,265]. Different *Borrelia* species could be preferentially associated with certain symptoms [266], which explains differences in disease manifestations observed between the two continents [193,201,267]. For this reason, when reading scientific articles, it is important to consider the area in which the study has been carried out.

As opposed to other gram negative bacteria, *Borrelia* do not contain lipopolysaccharides on their surface [191]. They instead express a large number of different outer surface proteins, some of which turned out to be very useful as antigens in serological tests.

The rather special genome of *Borrelia* is composed of one linear chromosome and several linear and circular plasmids [191,268] (Figure 4). *Borrelia* strains are very heterogeneous, which is mainly due to differences in plasmid content [191]. The majority of lipoproteins important for adaptation to their different environments are encoded by genes located on these plasmids [268,269]. When *Borrelia* are kept in culture, loss of plasmids encoding genes important for infection need to be considered [191,268] and complicated the isolation of antigens. In addition to that, *Borrelia* express different proteins under *in vitro* conditions than in the vertebrate host or in the tick environment. For this reason, ELISA assays often contain VlsE on top of *Borrelia* whole cell extracts [203].

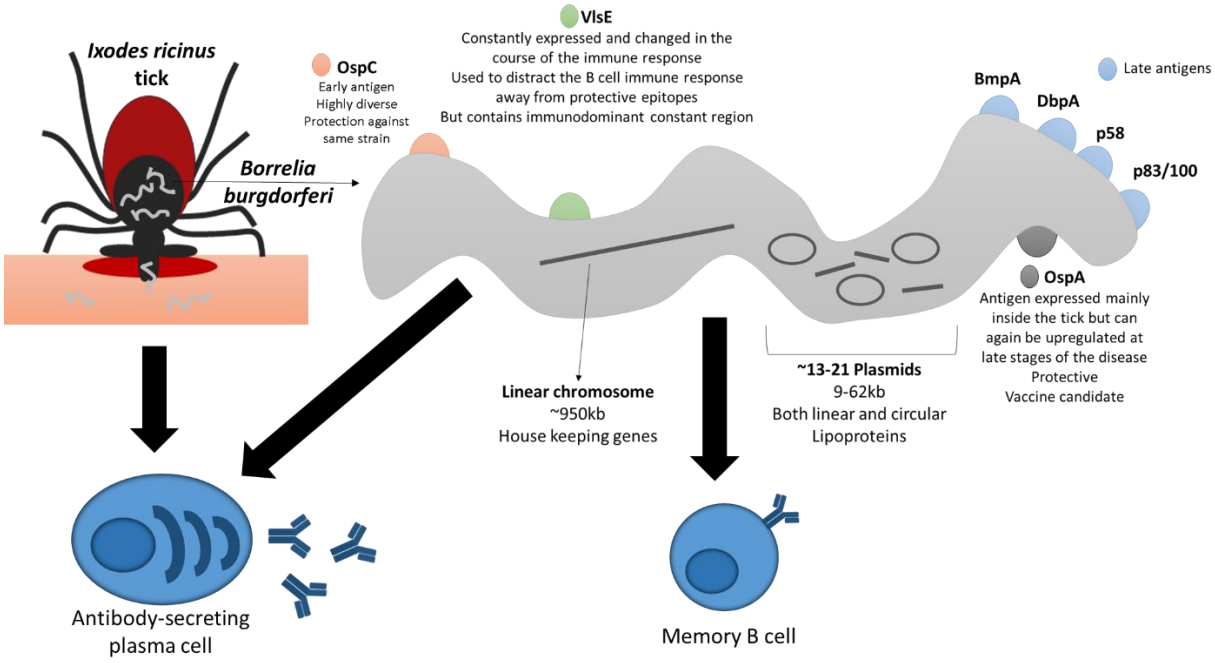


Figure 4: Both the tick vector but also *Borrelia* themselves actively influence the B cell immune response of mice and probably also of humans. *While the tick seems to inhibit only the local production of antibodies, Borrelia more extensively manipulate the B cell immune response in order to survive in their reservoir host. On one hand the great diversity of antigens (OspC) which are even actively changed in the course of the immune response (VlsE) prevents the host from being protected against different Borrelia strains. On the other hand Borrelia also directly affect the formation of long-lived plasma and memory B cells. Although the human immune response is also special – at least in some individual cases – it remains to be investigated more thoroughly what is going on at the cellular level in this case. The majority of housekeeping genes are expressed on the rather conserved linear chromosome of Borrelia, while many of the lipoproteins that are important for their transmission are expressed on the rather heterogeneous plasmids.*

Outer surface protein C (OspC) is a very important antigen used to support diagnosis. It is encoded on a circular plasmid and is one of the most diverse *Borrelia* proteins [203,268]. OspC is upregulated upon transition from the tick to the vertebrate host and is essential to establish an infection

[191,268,269]. Mouse experiments demonstrated, that the strong early antibody response induced by this diverse antigen is protective against the same strain but not strains expressing a different variant [270,271]. In humans, certain OspC types could be associated more frequently with disseminated disease [272,273].

In the mammalian host, OspC is only expressed transiently at early stages of infection [269], while expression of another very interesting antigen - variable major protein-like sequence expressed (VlsE) - is thought to increase over time [274]. This protein is expressed on a linear plasmid [191,268] and it was hypothesized that it is used by *Borrelia* to distract the B cell immune response away from protective epitopes [274]. During infection, *Borrelia* use genetic recombination to constantly change the sequence of this protein [274]. It was estimated, that with the used system, *Borrelia* could potentially generate 10^{30} different protein variants [275], a diversity that exceeds by far the expected diversity of the human B cell repertoire [276]. Paradoxically though, in its inner core, this protein contains a highly conserved epitope – invariable region 6 (IR6) - that is hidden in intact *Borrelia*, but which seems to be largely exposed to the human immune system *in vivo* [276]. Since antibodies elicited towards this region cannot kill the bacteria, it was proposed that *Borrelia* use reactivity to this epitope – in addition to the variable regions on intact bacteria - as decoy to even further distract the immune response away from functionally more important epitopes [276]. IgM antibodies are preferentially generated to the intact VlsE protein that should be expressed on the surface of intact bacteria, while antibodies of the IgG isotype are generally generated against the C6 epitope, which is rather expected to be exposed in non-intact bacteria [277]. Although quite speculative, findings from this study might indicate that live *Borrelia* influence the IgM switching especially at early stages of the infection, while dead bacteria might no longer do that. Even if the mouse B cell repertoire might use different mechanisms to recognize this epitope region [278], animal experiments [279] allow to more precisely compare antibody cross-reactivities between different strains. In humans the infecting bacteria is often hard to isolate [173]. These experiments already indicated, that also with VlsE-C6, the outcome of ELISA results can be influenced by the combination of *Borrelia* strains used in the ELISA and those for infection [279]. Even though the VlsE-C6 region seemed quite well conserved, human serum panels from different regions showed differences in reactivities towards individual peptide epitopes from different strains, indicating that a mixture of epitopes from different strains is needed to detect all seropositive patient sera [280-283]. Although the IgG subclass distribution was not affected, patients with more disseminated and chronic disease manifestations were more likely to also develop antibodies against the membrane-proximal region of VlsE, indicating the appearance of

new epitopes in the course of prolonged infection [104,284]. Similar to the antibodies generated towards the IR6 region, also these antibodies were found to be unable to bind to intact bacteria [284].

While antibodies against OspC and VlsE-C6 appear already at very early stages of infection, those towards DbpA, BmpA, p58 and p83/100 – which are also frequently used in immunoblots - only appear later [285]. The presence of specific bands can be useful to support diagnosis of certain disease conditions (stages), their absence however should be interpreted with caution. A negative result can also mean that the infecting strain expressed a different variant that was not detected by the assay.

The described antigenic proteins are just a few very well studied examples that allow to illustrate the complexity of the B cell immune response towards *Borrelia*. A much larger number of antigens is expected to be expressed inside the human host and many of them might still be unknown. As already mentioned before, because *Borrelia* express different proteins in different environments, it was very hard to isolate the antigens that are expressed *in vivo*. Poljak et al. chopped the genome of *Borrelia* into little pieces and expressed the corresponding amino acid sequences on the surface of bacteria [286]. Isolation of bacterial clones from these libraries with human *Borrelia*-reactive sera led to the identification of 122 different *Borrelia afzelii* antigenic regions [286]. Since these libraries are probably missing many discontinuous epitopes and largely focus on only those regions that show a high reactivity in common between different individuals, one can assume that the real number of antigenic regions largely exceeds this number. Taking also into account, that different *Borrelia* strains might express different variants of the same proteins, one can conclude that the *Borrelia* reactive B cell repertoire is probably highly complex and diverse among different individuals and timepoints.

1.2.3 The B cell immune response in Lyme disease

1.2.3.1 Findings from the mouse model

Mouse experiments showed that both the tick [287,288] but also *Borrelia* [289-293] themselves actively influence the B cell immune response (Figure 4). The tick seems to inhibit the local production of antibodies secreted by plasma cells, not however the formation of memory B cells [287,288]. In order to be able to feed on their hosts for prolonged times, ticks have developed various

different mechanisms to interfere with host immune responses [177-180]. This could explain, why the B cell immune response to tick-transmitted *Borrelia* appears delayed as compared to that one against tissue-transplanted bacteria [290]. A B cell inhibitory protein could also be isolated from tick salivary glands extracts [294].

It could be shown that *Borrelia* actively migrate to the draining lymph nodes and influence germinal center responses [290]. Large expansions of extrafollicular B cells could be observed [290]. Many of these B cells are reacting with *Borrelia* antigens, indicating a specific response [290]. Although with a delay, normal germinal centers are initiated at early stages of infection [289]. They however deteriorate later on followed by an accumulation of mainly naive B cells [289,293]. Lymph node architectural changes seem to already occur prior to the accumulation of B cells, meaning that the excessive accumulation of cells is not the reason for the observed deterioration[293]. The architectural changes were shown to be accompanied by an only transient increase in antibody avidity, indicating a failure to produce robust protective antibodies [291]. For many months after the infection, the mice are not able to generate a robust memory response [292]. Although the excessive accumulation of B cells in the lymph nodes appeared rather T cell independent, the formation of antibody secreting plasma cells and their long-term survival in bone marrow niches that occurred only at very late stages of infection were affected in the absence of T cell help [289]. In the absence of T helper cells, *Borrelia* burden is increased in mice, indicating that these cells affect the quality of the response [291]. Ectopic germinal centers are thought to be responsible for the occurrence of the delayed long-lived plasma cells, however their existence still needs to be proven [289]. The B cell immune response to *Borrelia* in this model was largely predominated by cells expressing the IgM isotype, indicating that they might have problems to class-switch [289]. This is in line with human studies, that showed an unusually long persistent *Borrelia*-positive IgM antibody titer in a subset of patients [295].

Mice infected with *Borrelia* also failed to induce a robust protective and long-term immune response to a co-administered influenza vaccine [292]. This confirms a direct effect of *Borrelia* on the B cell immune response. The exact mechanism these bacteria use to do that is however not totally clear yet. T cells of infected mice seem to follow a normal activation and differentiation route without the appearance of exhausted or suppressive cells [291]. T cells from infected donor mice however induced more B cells to express BLIMP-1 [291], which is known to be important for conferring B

cells to the plasma cell fate [296]. Congruently, they induced B cells to prematurely exit the proliferation phase and to more readily differentiate into antibody producing plasma cells [291]. Interestingly, this effect was antigen dependent, which is in line with the antigen-specificity of the extrafollicular B cell expansion described before [290,291]. It seems, that *Borrelia* try to prevent B cells from entering germinal center responses in which highly specific antibodies are usually generated. Like this, the body needs to protect himself with rather more polyspecific and suboptimal naïve B cell receptors of IgM isotype that are taken directly from the primary repertoire.

Besides the described effects on T cell help, it was observed that the follicular dendritic cell network [292] is disrupted and it seems that MyD88- and TRIF-independent type I IFN signalling [293] is involved but not the only factor in the excessive accumulation of B cells in the lymph nodes. In the absence of type I IFN signalling, mice only showed a reduced B cell accumulation in lymph nodes, but this did not prevent disorganization of germinal centers upon *Borrelia* infection [293]. Interestingly in humans it was found, that strains with a higher tendency to disseminate induce higher levels of type I and type III interferons [297]. Since *Borrelia* can exert a strong mitogenic activity on B lymphocytes [298], one initially thought that this might be at the origin of the extensive expansion of B cells. Toll-like receptor signalling was however excluded as the cause of this effect [293,299]. Since *Borrelia* are known to affect the complement system which also plays a role in B cell activation, it might be worth investigating its role in future projects [98,300,301].

Borrelia use two strategies to be transmitted from tick to tick. Either they are directly transferred to neighbour ticks that are sucking blood at the same time on the same animal (a process that is also known as "co-feeding") [302]. In other cases, they manage to evade the host immune response thereby ensuring persistence until the next tick is feeding. Rodents belong to this class of natural reservoir hosts for *Borrelia*, meaning that the bacteria can persist without causing too much damage to the animal [303]. Incidental hosts like humans are not as well adapted to *Borrelia* [303], and it was found that in Northern Europe for example 40% of infected individuals develop symptoms upon infection [192]. This number might vary though depending on circulating *Borrelia* strains in the region of interest. One needs to be cautious when transferring conclusions gained from mouse experiments to the human situation. The used mouse model is however mimicking many phenomena also observed in patients and also human Lyme disease is associated with a rather unusual B cell immune response.

1.2.3.2 The situation in humans

Early studies followed human antibody kinetics in the course of *Borrelia* infection and demonstrated a rather heterogeneous and slow response, which correlates with the duration of symptoms and spirochete dissemination prior to the start of antibiotic therapy [295,304-306]. It was found, that treatment with antibiotics largely affects the evolution of the B cell immune response to Lyme disease. Patients under treatment could be subdivided into different reactivity groups. Some developed persistent antibody titers, while others stayed seronegative and even others showed a decline in antibody titers over time [304]. Some patients had problems to induce robust long-lived IgG antibody responses even when antibiotics treatment was only started several months after onset of symptoms [304]. Discrepant results between different serological test systems [209,211,280,307] however make it difficult to draw general conclusions about the presence or absence of a B cell immune response in a certain patient. This indicates however, that differences in infecting *Borrelia* strains very likely affect the outcome of the B cell immune response. What can be concluded is, that the B cell immune response to *Borrelia* is highly diverse among patients. On one hand the infecting *Borrelia* strain but on the other hand also host genetic factors might contribute to the frequently observed discrepant results among different test systems.

A decline in antibodies after antibiotics treatment is very frequently observed in Lyme disease [305,308,309]. Interestingly, a recent study showed more discrepant results among serological tests in treated as opposed to healthy or acute donors [212], implying that antibiotics might not affect all detected antibodies to the same extent. After treatment, many early stage patients remain seronegative [305]. Although it was not possible to test whether these individuals would have seroconverted at later timepoints, current optimized versions of serological tests indicate that the majority of patients generate detectable antibodies especially at later stages of the disease [173,206,207]. Due to the presence of a higher number of different *Borrelia* genospecies [265], the situation is more complicated in Europe than in the United States [310]. Since so many different *Borrelia* strains can cause infections, a seronegative result does not necessarily mean that there is nothing going on at the B cell level, but the used test might not contain the right *Borrelia* antigens or was simply not sensitive enough to detect a response [210].

Some patients were shown to develop a long lasting IgM antibody titer that stayed detectable for up to 20 years after the infection [295]. This is rather unusual, as in other diseases antibodies of this isotype are only present at acute stages which is why they are frequently used for diagnosis [5]. Reinfection of early stage patients from endemic areas [219,220,311] is indicating a weak non-protective antibody response and might further complicate the interpretation of serological tests. It needs to my knowledge however still be investigated more thoroughly how pre-existing antibodies influence the chance of getting reinfected. Absence of protection can be explained by different means. Similar to what has been found in the mouse model, *Borrelia* might actively influence the human B cell immune response and prevent the induction of protective antibodies especially at early stages of the disease. In accordance with that, sera from late stage patients more likely protected mice from infection than that of early stage subjects [312,313]. Although also other antibodies might have been involved, protection was largely attributed to antibodies generated against outer surface proteins (OspA and OspB) [312,313]. OspA is expressed inside the tick and downregulated as *Borrelia* enter the animal host. This protein can occur at the surface of bacteria again at later stages of infection, which is the reason why some patients generated these protective antibodies. In this example protection was dependent on the expression of the protein rather than the body not being able to generate antibodies. OspA ended up in a licensed vaccine, which was however withdrawn from the market due to low efficacy combined with complaints about adverse side effects [186]. New vaccination strategies based on this protein and a combination with OspC epitopes – a protein that is expressed *in vivo* - are current research topics [314-316]. Inclusion of OspA from only one *Borrelia* strain was not enough to elicit protection against all other circulating species. This means, that protection might also depend on the protein variants expressed by the infecting strains. In agreement with this are statistical analysis performed on ticks and erythema migrans patients from the same area, which indicate that even early stage patients might be protected for several years against the same but not a different *Borrelia* strain [221]. This assumption is based on the finding, that strains isolated from first infections are generally different from those of subsequent infections [220]. Mouse studies indicate that reinfection with the same *Borrelia* strain might however also be possible especially after antibiotics treatment [292]. To conclude, some patients might indeed be protected against certain *Borrelia* strains. However this might depend on which strains they catch (and caught before) as well as the list and nature of antigens expressed prior to their elimination. It seems, that the outcome of the B cell immune response is influenced by a complex interplay between host genetics and virulence factors expressed by the *Borrelia* strain.

Although many human studies focus on the antibody response, to my knowledge only a few studies have investigated the human B cell immune response to *Borrelia* at a phenotypic level. In contrast to some older studies that showed an increase in peripheral B-lymphocytes in other acute bacterial infections [142,143], Lyme disease does not seem to be associated with such changes [317,318]. Interestingly, in Neuroborreliosis [230,319,320] but also in Lyme arthritis [240,321,322], tissue localized ectopic *Borrelia* specific B cell immune responses have been observed. In the case of Neuroborreliosis, those were associated with low numbers of *Borrelia*-specific B cells in peripheral blood [319]. The strong response of synovial mononuclear cells from Lyme arthritis patients confirms a local immune reaction occurring within the joints [323].

A recent RNAseq study of acute patient PBMCs found differences in the induction of B cell developmental pathways and calcium-induced T cell apoptosis by *Borrelia* as compared to other infectious agents [251]. Furthermore their data generated evidence for the involvement of a rather low proportion of B and T cells in peripheral blood in the acute Lyme disease condition as compared to other diseases. Although it still needs to be tested whether ectopic B cell immune responses are responsible for the effects observed in the aforementioned mouse model, their occurrence can be considered as rather confirmed in the human setting. The not necessarily but often transient nature of ectopic lymphoid follicles (ELFs) [324] would explain the patterns of reduction in antibody titers observed after antibiotics treatment [305,308,309]. The general absence of those structures from the skin[324], could explain the delayed antibody response that seems to correlate with the degree of *Borrelia* dissemination [308,309,323]. Recent studies that were able to associate Th17 responses [223,325] with post-Lyme disease symptoms and the development of autoantibodies at later stages of the disease, further support the idea, that inflammation driven ectopic B cell immune responses could play a role in the chronic course observed in a subset of patients. Th17 cells seem to play a role in ectopic lymphoid follicle immune reactions in chronic inflammation and these structures have been shown to be a possible trigger in many autoimmune and chronic diseases [324]. In this context, the findings from the mouse study that showed, that chronic inflammatory conditions can prevent the homing of long-lived plasma cells to the bone marrow is also important to consider [154].

Analysis of immune factor levels within peripheral blood was able to separate acute patients into two groups, one with higher levels of T cell recruiting chemokine and inflammatory marker expression associated with lymphopenia and the development of antibodies against *Borrelia* and a second group

in which B and T cell responses seem to be rather low [248], indicating heterogeneous responses among patients. This is in line with the rather heterogeneous nature also observed at the antibody level [295,304,305].

1.3 Objectives of this study

As outlined above, the human B cell immune response to *Borrelia* is very complex and far from being completely understood. Especially studies at the B cell receptor and cellular levels are missing. The main goal of this project was to explore whether signatures could be extracted from the bulk B cell repertoire of acute Lyme disease patients that might be useful to support diagnosis. Since B cells generally start to react with antigens already very early after contact [53], we expect to be able to detect an ongoing response at the cellular level already way before the antibody response is measurable. This is especially interesting in this case, as it was shown that ticks might inhibit the production of antibodies, but they don't seem to interfere with the generation of memory B cells [287,288].

Since B cells are the cells of our body that react most directly and specifically to immune challenges, we expect similar clones to appear in different individuals. Even if a very large amount of different B cell clones might be involved in the response, this does not exclude the possibility that some clones might nevertheless be generated in common. It is very likely though, that they only occur at very low levels. Recent studies indicate, that some degree of convergence is occurring [95,127,165,169,171,326], which looks highly promising and supports this hypothesis. One big problem though is the great diversity of the B cell repertoire. Even with next generation sequencing, we are very likely missing a large portion of important B cell clones [122]. For this reason, focusing on antigen-specific sequences is crucial.

The IR6 domain of VlsE (VlsE-C6) is an epitope region, that is quite well conserved between different *Borrelia* strains [276]. Antibodies are already detectable very early and throughout all stages of infection probably as a result of it being expressed throughout the whole infection period [274,327]. For this reason, we decided to use this epitope as model antigen to test for the level and kinetics of *Borrelia* reactive B cells in peripheral blood of acute Lyme disease patients.

Besides the potential presence of immune signatures, B cell repertoire next-generation sequencing data harbour a tremendous amount of different interesting information that might be valuable to better understand the human B cell immune responses towards pathogens. In addition to that, we might be

able to confirm some of the findings from the mouse model which clearly showed that *Borrelia* actively influence the B cell immune response [289-293]. We expect, that certain features (like mutation status, class-switching of clones, evolution of clones over different timepoints,…) might allow us to draw conclusions about the nature of the B cell immune response towards *Borrelia* and maybe even to separate different patients into reactivity groups. In the case of hepatitis C virus infection, differences in clonal evolutions could for example be found between resolving and chronically evolving patients [328]. It would be highly interesting to determine whether similar phenomena are also occurring in response to Lyme disease infection for which a chronic course does also exist.

2 MATERIALS

2.1 Chemicals and solutions used to amplify and purify nucleic acids

Compound	Company
10x Dulbecco's Phosphate buffered saline (DPBS) w/o Ca^{++}, Mg^{++}	Lonza
1kb plus DNA ladder	Thermo Fisher Scientific (Invitrogen)
2-Propanol BioReagent for molecular biology	Sigma-Aldrich
Dimethyl sulfoxide, Hybri-Max™, sterile-filtered, BioReagent, suitable for hybridoma, 99.7% (DMSO)	Sigma-Aldrich
Dithiothreitol (DTT) (0.1M)	Thermo Fisher Scientific (Invitrogen)
Ethanol molecular biology grade	VWR
IGEPAL® CA-630 for molecular biology (Nonidet P-40/Igepal)	Sigma-Aldrich
Magnesium Chloride ($MgCl_2$)	Thermo Fisher Scientific (Invitrogen)
Nucleotides (dNTPs)	Thermo Fisher Scientific (Invitrogen)
Oligonucleotides/Primers	Eurogentec
Orange DNA Loading Dye (6x)	Thermo Fisher Scientific
Random primers	Thermo Fisher Scientific (Invitrogen)
RNAse/DNase free water	Thermo Fisher Scientific (Gibco)
SeaKem® LE Agarose	Lonza
SYBR®Safe DNA Gel Stain	Thermo Fisher Scientific (Invitrogen)

2.2 Chemicals used to prepare buffers, solutions and media

Compound	Company
10x Dulbecco's Phosphate buffered saline (DPBS) w/o Ca^{++}, Mg^{++}	Lonza
2-Amino-2-methyle-1-propanole (AMP)	Sigma-Aldrich
Bovine Serum Albumin	Sigma-Aldrich
Carbonate-Bicarbonate Buffer capsules	Sigma-Aldrich
Dimethyl sulfoxide, Hybri-Max™, sterile-filtered, BioReagent, suitable for hybridoma, 99.7% (DMSO)	Sigma-Aldrich
Ethylenediaminetetraacetic acid (EDTA)	Sigma-Aldrich
Fetal bovine serum (FBS), South American Origin, Low IgG, sterile filtered, for cell culture (Lot: 8SB0001):	Lonza
Ficoll-Paque™ PLUS, endotoxin tested, sterile	GE Healthcare Life Sciences
HyClone™ Fetal Bovine Serum (U.S.), Standard (Lot: AAB199095)	GE Healthcare Life Sciences
$MgCl_2.6H_2O$	Sigma-Aldrich
Penicillin-Streptomycin (Pen/Strep) (10000U/ml)	Thermo Fisher Scientific (Gibco)
Phosphatase substrate 5mg tablets	Sigma-Aldrich
Potassium Chloride (KCl)	Sigma-Aldrich
RPMI 1640 with UltraGlutamine I and HEPES	Lonza
Sodium Acetate (CH_3COONa)	Merck Millipore
Sodium azide (NaN3)	Sigma-Aldrich
Sodium Chloride (NaCl)	Acros Organics
Sodium pyruvate (100mM)	Thermo Fisher Scientific (Gibco)
Tris(hydroxymethyl)aminomethane (Tris)	Sigma-Aldrich

Trizma® Acetate	Sigma-Aldrich
Trizma® base	Sigma-Aldrich
TWEEN® 20	Sigma-Aldrich
UltraGlutamine I	Lonza

From these reagents, the following buffers were prepared:

Buffer / solution	**Reagent**	**Volume / Concentration**	**Experiment**
TAE buffer (50x, pH 7.8)	Tris Sodium Acetate EDTA	2M 25mM 0.5M	*Agarose gel electrophoresis*
Freezing solution 1 (complete RPMI1640 medium, 4°C)	RPMI 1640 Sodium pyruvate Pen/Strep FBS (Lonza)	1x 1% 1% 5%	*PBMC cryopreservation*
Freezing solution 2 (4°C)	FBS (Lonza) DMSO	1x 20%	*PBMC cryopreservation*
FACS buffer (4°C)	PBS NaN_3 FBS (Lonza)	1x 0.1% 2%	*Multicolor flow cytometry*
FACS buffer (4°C)	PBS FBS (Lonza)	1x 2%	*Tetramer staining*
MACS buffer (4°C)	PBS FBS EDTA	1x 0.5% 2mM	*B cell isolation with Miltenyi MACS kit*
Carbonate buffer (4°C)	Carbonate-Bicarbonate capsule Bidi water	1x 100ml	*ELISA*
Washing buffer (RT)	TWEEN® 20 Sodium Chloride Trizma® base	1% 154mM 10mM	*ELISA*
Dilution buffer (RT)	TWEEN® 20 BSA Trizma® Acetate Sodium Chloride Potassium Chloride	0.1% 1% (w/v) 15mM 136mM 2mM	*ELISA*
Blocking buffer	BSA Trizma-Acetate Sodium Chloride	1% (w/v) 15mM 136mM	*ELISA*

	Potassium Chloride	2mM	
Substrate buffer (pH 10.2)	2-Amino-2-methyle-1-propanole $MgCl_2.6H_2O$	1mM 0.1mM	*ELISA*
Washing buffer	PBS TWEEN® 20	1x 0.05%	*384 ELISA*
Blocking buffer	PBS TWEEN® 20 BSA	1x 0.01% 10%	*384 ELISA*

2.3 Enzymes

Enzyme	**Company**
HotStarTaq® DNA Polymerase	Qiagen
Phusion® High-Fidelity DNA Polymerase	New England Biolabs
Platinum® Taq DNA polymerase	Thermo Fisher Scientific (Invitrogen)
Q5® Hot Start High-Fidelity DNA Polymerase	New England Biolabs
RNaseOUT™ Recombinant Ribonuclease Inhibitor	Thermo Fisher Scientific (Invitrogen)
RNasin® Ribonuclease Inhibitor	Promega
SuperScript® III Reverse Transcriptase	Thermo Fisher Scientific (Invitrogen)

2.4 Antibodies, and reagents used for flow cytometry

Antibody / Chemical	**Company**
Goat Anti-Human IgG-AP	Southern Biotech
Hoechst	Thermo Fisher Scientific (Invitrogen)
Anti-Human IgA-APC	Jackson ImmunoResearch Laboratories
Anti-Human IgM-APC	Jackson ImmunoResearch Laboratories
Mouse Anti-Human CD10-BV510	BD Horizon

Mouse Anti-Human CD138-APC	Biolegend
Mouse Anti-Human CD138-BV711	BD Horizon
Mouse Anti-Human CD14-eFluor605NC	eBioscience
Mouse Anti-Human CD14-FITC	Immunotools
Mouse Anti-Human CD14-PE	Immunotools
Mouse Anti-Human CD16-PercPcy5.5	Biolegend
Mouse Anti-Human CD19-BV605	BD Pharmingen
Mouse Anti-Human CD20-AF488	Biolegend
Mouse Anti-Human CD20-AF700	Biolegend
Mouse Anti-Human CD20-Biotin	Immunotools
Mouse Anti-Human CD21-PE-Cy7	Biolegend
Mouse Anti-Human CD23-APC-eFluor780	eBioscience
Mouse Anti-Human CD24-BV421	BD Horizon
Mouse Anti-Human CD24-eF450	eBioscience
Mouse Anti-Human CD27-BV421	BD Biosciences
Mouse Anti-Human CD27-PECF594	BD Horizon
Mouse Anti-Human CD38-PerCP-Cy5.5	Biolegend
Mouse Anti-Human CD3-FITC	Immunotools
Mouse Anti-Human CD3-PE	Immunotools
Mouse Anti-Human CD3-PE-Dy647	Immunotools
Mouse Anti-Human CD43-APC	eBioscience
Mouse Anti-Human CD56-PeCP-Cy5.5	Biolegend
Mouse Anti-Human CD5-FITC	Immunotools
Mouse Anti-Human IgD-BV421	BD Pharmingen
Mouse Anti-Human IgD-PacBlue	Biolegend
Mouse Anti-Human IgD-PE	BD Biosciences

Mouse Anti-Human IgG-PE	BD Pharmingen
Mouse Anti-Human IgM-BV570	Biolegend
Neutravidin-DyLight650	Thermo Fisher Scientific

2.5 Commercial kits

Kit	**Company**
Agencourt® AMPure® XP beads	Beckman Coulter
Agilent High Sensitivity DNA Kit	Agilent Genomics
AllPrep DNA/RNA Mini Kit	Qiagen
Big Dye® Terminator v3.1 Cycle Sequencing Kit	Thermo Fisher Scientific (Applied Biosystems)
Borrelia afzelii + VlsE IgG Europe ELISA Testkit	Genzyme Virotech
Borrelia afzelii IgM ELISA Testkit	Genzyme Virotech
Borrelia Europe LINE IgM Immunoblot	Genzyme Virotech
Borrelia Europe Plus TpN17 LINE IgG Immunoblot	Genzyme Virotech
Dynabeads™ mRNA DIRECT™ Purification Kit	Thermo Fisher Scientific (Invitrogen)
EasySep™ Human B cell Enrichment Kit	Stemcell Technologies
Human B cell Isolation Kit II	Miltenyi Biotec
Human IgG total Ready-SET-Go!® ELISA	eBioscience
Ion 318™ Chip Kit v2	Thermo Fisher Scientific
Ion PGM™ Enrichment Beads	Thermo Fisher Scientific
Ion PGM™ Wash 2 Bottle Kit	Thermo Fisher Scientific
Ion PGM™ Hi-Q™ Sequencing Kit	Thermo Fisher Scientific
Ion PGM™ Sequencing 400 Kit	Thermo Fisher Scientific

Ion PGM™ Template OT2 400 Kit	Thermo Fisher Scientific
Ion Sphere™ Quality Control Kit	Thermo Fisher Scientific
Jet Quick PCR Purification Spin kit	Genomed
QIAamp® DNA Blood Mini kit	Qiagen
QIAquick® Gel Extraction kit	Qiagen

2.6 Primers

2.6.1 Primers used for amplification of *Borrelia* DNA

Primer	5'-3' Sequence	PCR reaction	Reference
V1a	GGGAATAGGTCTAATATTAGC	*First round PCR*	[329]
V1b	GGGGATAGGTCTAATATTAGC	*First round PCR*	[329]
V3a	GCCTTAATAGCATGTAAGC	*Second round PCR*	[329]
V3b	GCCTTAATAGCATGCAAGC	*Second round PCR*	[329]
R1	CATAAATTCTCCTTATTTTAAAGC	*First & second round PCRs*	[329]
R37	CCTTATTTTAAAGCGGC	*First & second round PCRs*	[329]

2.6.2 Primers used for amplification of *Rickettsia* DNA

Primer	5'-3' Sequence	PCR reaction	Reference
Rr17k.1p	TTTACAAAATTCTAAAAACCAT	*First round PCR*	[330]
Rr17k.539n	TCAATTCACAACTTGCCATT	*First round PCR*	[330]
Rr17k.90p	GCTCTTGCAACTTCTATGTT	*Second round PCR*	[330]
Rr17k.539n	TCAATTCACAACTTGCCATT	*Second round PCR*	[330]

2.6.3 Primers used for amplification of immunoglobulin heavy chain genes from single cells

Primer	5'-3' Sequence	PCR reaction	Reference
VHL-1	TCACCATGGACTGSACCTGGA	*First round PCR*	[331]
VHL-2	CCATGGACACACTTTGYTCCAC	*First round PCR*	[331]
VHL-3	TCACCATGGAGTTTGGGCTGAGC	*First round PCR*	[331]
VHL-4	AGAACATGAAACAYCTGTGGTTCTT	*First round PCR*	[331]
VHL-5	ATGGGGTCAACCGCCATCCT	*First round PCR*	[331]
VHL-6	ACAATGTCTGTCTCCTTCCTCAT	*First round PCR*	[331]
CμII	CAGGAGACGAGGGGGAAAAG	*First round PCR*	[331]
CγII	GCCAGGGGGAAGACSGATG	*First round PCR*	[331]
CαII	GCTCAGCGGGAAGACCTT	*First round PCR*	[331]
VH-1-nr	CAGGTSCAGCTGGTRCAGTC	*Second round PCR*	[331,332]
VH-2-nr	CAGRTCACCTTGAAGGAGTC	*Second round PCR*	[331,332]
VH-3-nr	SAGGTGCAGCTGGTGGAGTC	*Second round PCR*	[331,332]
VH-4-nr	CAGGTGCAGCTGCAGGAGTC	*Second round PCR*	[331,332]
VH-5-nr	GARGTGCAGCTGGTGCAGTC	*Second round PCR*	[331,332]
VH-6-nr	CAGGTACAGCTGCAGCAGTC	*Second round PCR*	[331,332]
CμIII-nr	GAAAAGGGTTGGGGCGGATGC	*Second round PCR*	[331,332]
CγIII-nr	GACSGATGGGCCCTTGGTGGA	*Second round PCR*	[331,332]
CαIII-nr	GACCTTGGGGCTGGTCGGGGA	*Second round PCR*	[331,332]

2.6.4 Primers used for amplification of immunoglobulin light chain genes from single cells

Primer	5'-3' Sequence	PCR reaction	Reference
VκL-1	GCTCAGCTCCTGGGGCTCCTG	*First round PCR*	[331]

VκL-2	CTGGGGCTGCTAATGCTCTGG	*First round PCR*	[331]
VκL-3	TTCCTCCTGCTACTCTGGCTC	*First round PCR*	[331]
VκL-4	CAGACCCAGGTCTTCATTTCT	*First round PCR*	[331]
VλL-1	CCTCTCCTCCTCACCCTCCT	*First round PCR*	[331]
VλL-2	CTCCTCACTCAGGGCACA	*First round PCR*	[331]
VλL-3	ATGGCCTGGAYCSCTCTCC	*First round PCR*	[331]
CκII	TTTCAACTGCTCATCAGATGGCGG	*First round PCR*	[331]
CλII	AGCTCCTCAGAGGAGGGYGG	*First round PCR*	[331]
Vκ-1-nr	CGMCATCCRGWTGACCCAGT	*Second round κ PCR*	[331,332]
Vκ-2-nr	CGATRTTGTGATGACYCAG	*Second round κ PCR*	[331,332]
Vκ-3-nr	CGAAATWGTGWTGACRCAGTCT	*Second round κ PCR*	[331,332]
Vκ-4-nr	CGACATCGTGATGACCCAGT	*Second round κ PCR*	[331,332]
CκIII-nr	AAGATGAAGACAGATGGTGC	*Second round κ PCR*	[331,332]
AgeI-Vλ-1	CTGCTACCGGTTCCTGGGCCCAGTC TGTGCTGACKCAG	*Second round λ PCR*	[333,334]
AgeI-Vλ-2	CTGCTACCGGTTCCTGGGCCCAGTC TGCCCTGACTCAG	*Second round λ PCR*	[333,334]
AgeI-Vλ-3	CTGCTACCGGTTCTGTGACCTCCTAT GAGCTGACWCAG	*Second round λ PCR*	[333,334]
AgeI-Vλ-4/5	CTGCTACCGGTTCTCTCTCSCAGCYT GTGCTGACTCA	*Second round λ PCR*	[333,334]
AgeI-Vλ-6	CTGCTACCGGTTCTTGGGCCAATTTT ATGCTGACTCAG	*Second round λ PCR*	[333,334]
AgeI-Vλ-7/8	CTGCTACCGGTTCCAATTCYCAGRC TGTGGTGACYCAG	*Second round λ PCR*	[333,334]
XhoI-Cλ	CTCCTCACTCGAGGGYGGGAACAGA GTG	*Second round λ PCR*	[333,334]

2.6.5 Primers used for library preparation

Primer	5'-3' Sequence	PCR reaction	Reference
A-MID9-8N-4-8N-G	GCGTGTCTCCGACTCAGTGAGCGGAACNNNNNNNNGACTNNNNNNNNAAGACCGATGGGCCCTTG	*Reverse transcription (IgG, MID9)*	[1]
A-MID15-8N-4-8N-G	GCGTGTCTCCGACTCAGTCTAGAGGTCNNNNNNNNGACTNNNNNNNNAAGACCGATGGGCCCTTG	*Reverse transcription (IgG, MID15)*	[1]
A-MID21-8N-4-8N-G	GCGTGTCTCCGACTCAGTCGCAATTACNNNNNNNNGACTNNNNNNNNAAGACCGATGGGCCCTTG	*Reverse transcription (IgG, MID21)*	[1]
A-MID22-8N-4-8N-G	GCGTGTCTCCGACTCAGTTCGAGACGCNNNNNNNNGACTNNNNNNNNAAGACCGATGGGCCCTTG	*Reverse transcription (IgG, MID22)*	[1]
A-MID23-8N-4-8N-G	GCGTGTCTCCGACTCAGTGCCACGAACNNNNNNNNGACTNNNNNNNNAAGACCGATGGGCCCTTG	*Reverse transcription (IgG, MID23)*	[1]
A-MID25-8N-4-8N-G	GCGTGTCTCCGACTCAGCCTGAGATACNNNNNNNNGACTNNNNNNNNAAGACCGATGGGCCCTTG	*Reverse transcription (IgG, MID25)*	[1]
A-MID9-8N-4-8N-M	GCGTGTCTCCGACTCAGTGAGCGGAACNNNNNNNNGACTNNNNNNNNGGGAATTCTCACAGGAGACG	*Reverse transcription (IgM, MID9)*	[1]
A-MID15-8N-4-8N-M	GCGTGTCTCCGACTCAGTCTAGAGGTCNNNNNNNNGACTNNNNNNNNGGGAATTCTCACAGGAGACG	*Reverse transcription (IgM, MID15)*	[1]
A-MID21-8N-4-8N-M	GCGTGTCTCCGACTCAGTCGCAATTACNNNNNNNNGACTNNNNNNNNGGGAATTCTCACAGGAGACG	*Reverse transcription (IgM, MID21)*	[1]
A-MID22-8N-4-8N-M	GCGTGTCTCCGACTCAGTTCGAGACGCNNNNNNNNGACTNNNNNNNNGGGAATTCTCACAGGAGACG	*Reverse transcription (IgM, MID22)*	[1]
A-MID23-8N-4-8N-M	GCGTGTCTCCGACTCAGTGCCACGAACNNNNNNNNGACTNNNNNNNNGGGAATTCTCACAGGAGACG	*Reverse transcription (IgM, MID23)*	[1]
A-MID25-8N-4-8N-M	GCGTGTCTCCGACTCAGCCTGAGATACNNNNNNNNGACTNNNNNNNNGGGAATTCTCACAGGAGACG	*Reverse transcription (IgM, MID25)*	[1]
P1-V1-FR2	CTATGGGCAGTCGGTGATCTGGGTGCGACAGGCCCCTGGACAA	*Second strand synthesis*	[335]

P1-V2-FR2	CTATGGGCAGTCGGTGATTGGATCCGTCAGCCCCCAGGGAAGG	*Second strand synthesis*	[335]
P1-V3-FR2	CTATGGGCAGTCGGTGATGGTCCGCCAGGCTCCAGGGAA	*Second strand synthesis*	[335]
P1-V4-FR2	CTATGGGCAGTCGGTGATTGGATCCGCCAGCCCCCAGGGAAGG	*Second strand synthesis*	[335]
P1-V5-FR2	CTATGGGCAGTCGGTGATGGGTGCGCCAGATGCCCGGGAAAGG	*Second strand synthesis*	[335]
P1-V6-FR2	CTATGGGCAGTCGGTGATTGGATCAGGCAGTCCCCATCGAGAG	*Second strand synthesis*	[335]
P1-V7-FR2	CTATGGGCAGTCGGTGATTTGGGTGCGACAGGCCCCTGGACAA	*Second strand synthesis*	[335]
amp_A	CCATCTCATCCCTGCGTGTCTCCGACTCAG	*Amplification PCR*	
amp_P1	CCTCTCTATGGGCAGTCGGTGAT	*Amplification PCR*	

2.7 Instruments

Instrument	**Company**
2100 Bioanalyzer	Agilent Technologies
ABI PRISM® 3130xl Genetic Analyzer	Applied Biosystems
CFX96™ Real-Time PCR Cycler	Bio-Rad
Electrophoresis Power Supply EV231	Consort
FACSAria SORP	BD Biosciences
Gel tank combs and casting form	Bioplastics
Heating block	Peqlab
InGenius gel documentation system	Syngene
Ion OneTouch™ 2 System	Thermo Fisher Scientific
Ion PGM™ Sequencer	Thermo Fisher Scientific

Leica M205C Stereomicroscope equipped with camera and incident LED source	Leica
Leitz DMIL Inverted Phase Contrast Microscope	Leitz
NanoDrop ND-1000 Spectrophotometer	Thermo Fisher Scientific
Precision balance	Sartorius
Purelab® flex water purifier	ELGA LabWater
Qubit® 2.0 Fluorometer	Thermo Fisher Scientific (Invitrogen)
Safe Imager™ Tramsilluminator	Thermo Fisher Scientific (Invitrogen)
SL40R centrifuge	Thermo Scientific
SpectraMax Plus Microplate Reader	Molecular devices
Tissue Lyser II	Qiagen
UNO96 Thermal Cycler	VWR

2.8 Software and bioinformatics

Software / Tool	**Company / Developer**
2100 Expert Software	Agilent Genomics
bcRep R package (bcRep_1.3.4)	Bischof and Ibrahim [2]
BioEdit Sequence Alignment Editor v7.2.5.0	Tom Hall
Bio-Rad CFX Manager	Bio-Rad
Cd-hit-v4.6.7	Li et al. and Fu et al. [336,337]
Change-o	Gupta et al. [338]
Fastx-Toolkit	The Hannon Lab
GeneSnap	Syngene
GraphPad Prism 5	GraphPad Software, Inc.
IgBlast	Ye et al. [339]

IMGT®/HighV-QUEST	Alamyar et al. [340]
IMGT®/V-QUEST	Brochet et al. [341]
Kaluza Analysis Software	Beckman Coulter
PAGAN graph aligner	Löytynoja et al. [342]
Phylogeny.fr	Laboratoire d'Informatique, de Robotique et de Microélectronique de Montpellier [343-346]
Python 2.7.6	Python Software Foundation
Python clustering script (cl_1.1.9)	Galson et al. [126,127,171]
R	The R Foundation
Rstudio	Rstudio
SeqTrace-0.9.0	Stucky [347]
Torrent Suite Software and Torrent Server	Thermo Fisher Scientific
VDJtools-1.0.3	Shugay et al. [348]

2.9 Blood tubes and other special lab materials

Lab material	**Company**
Axygen® 0.2mL Maxymum Recovery® Thin Wall PCR Tubes	Axygen Scientific
BD Vacutainer® K2 EDTA tube (10ml)	Thermo Fisher Scientific
Cryo.s™ Cryogenic Storage Vials	Greiner Bio One
DynaMag™-2 Magnet	Thermo Fisher Scientific
EasySep™ Magnet	Stemcell Technologies
Eppendorf® DNA LoBind microcentrifuge tubes 1,5 ml	VWR
LeucoSep® tube, polypropylen, 50ml, sterile	Greiner Bio One
LS and MS columns	Miltenyi Biotec

MACS® Manual Separators (OctoMACS and QuadroMACS)	Miltenyi Biotec
Polyester films, non sterile	VWR
VACUETTE® RÖHRCHEN 8 ml Z Serum Separator	Greiner Bio One
VACUETTE® TUBE 9 ml NH Sodium Heparin	Greiner Bio One

3 METHODS

3.1 Recruitment of patients

For this study, patients for which Lyme borreliosis has been diagnosed by a medical doctor from Luxembourg have been recruited. A total of three blood draws were planned for each patient (Figure 5). In general, the first blood draw was done as close as possible to the timepoint of diagnosis (average duration of symptoms before first blood draw: 8 ±13 weeks ranging from 2 days to 1 year) and start of therapy (average before first blood draw: 4±5 days ranging from 0 to 17 days). A second blood draw was performed one week after and a third one, one month after the first visit. Due to the expected constant state of seropositive, seronegative and chronic Lyme disease patients, these donors were only sampled at a single timepoint. At each visit, the patients had to fill out a questionnaire asking them about their symptoms, diagnosis, treatment, tick exposure and previous Lyme disease infections (for more information, please refer to 8.1 Questionnaires). Only patients which reported at their last blood draw that antibiotics treatment reduced, changed or healed their symptoms or for which a clear picture of the erythema migrans was available were considered as clear acute cases and were included for the analysis of the B cell immune response against *Borrelia*. Since the overall number of participants was not high, we continued recruitment until the end of the study and therefore the data presented are not always consistently performed on the same patients. Since this study was rather exploratory in terms of getting first hints about the B cell immune response at the repertoire level, for subsequent analysis, we chose to study the most clear and interesting cases available at the time the assay was performed over the ones we started to explore already. Table 3, Table 4 and Table 5 give a general overview and more detailed information about relevant participants from this study. To keep groups comparable, age matched donors were used for each experiment. As control cohort, healthy individuals or donors with a recent tick bite were included. Ticks of these donors were screened for the presence of *Borrelia* and *Rickettsia* DNA [349]. Donors for which at least the back of the tick was still intact (adult female *I. ricinus* tick) and for which the *Borrelia* PCR was negative were included into the control group. These donors were selected based on their age to equalize acute and control groups age distribution. There were more female participants in the acute Lyme disease

group as compared to the controls. The current study has been accepted by Luxembourg's ethics committee and data protection commission and each donor had to sign an informed consent before blood was taken.

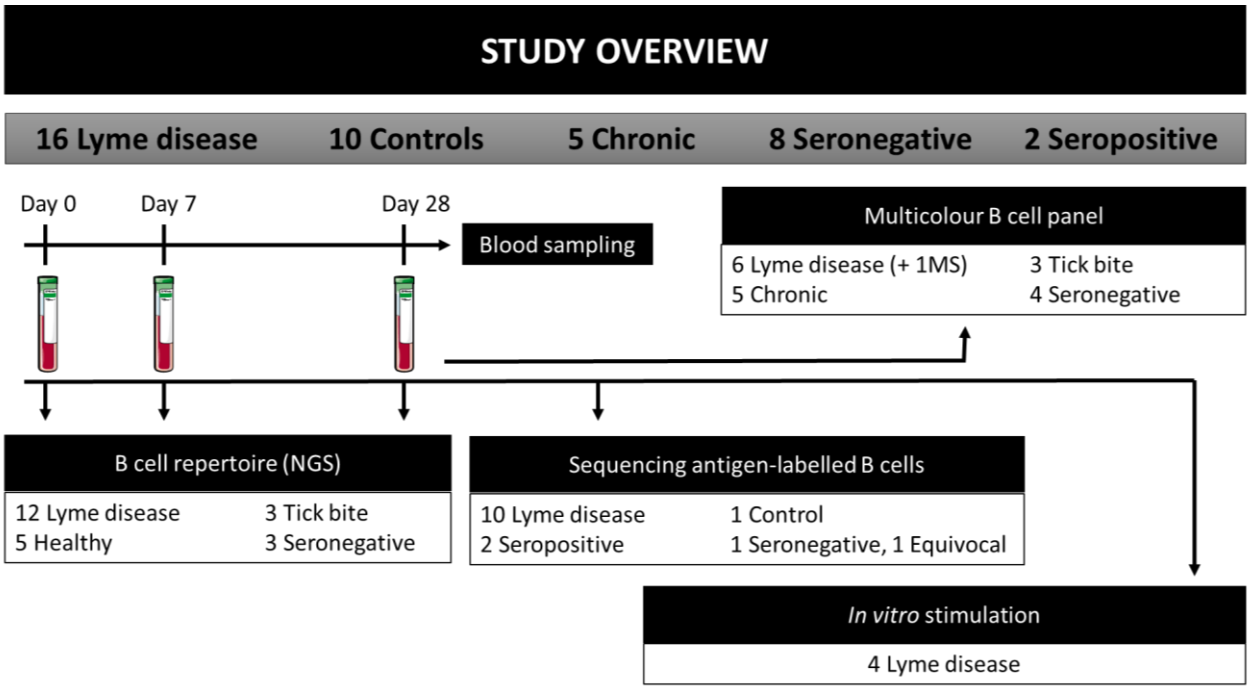


Figure 5: Overview of blood sampling and experiments performed on the samples.

3.2 Estimation of tick feeding time

To be able to estimate the time of feeding [350,351] and to assess intactness of the ticks after removal, a picture from the top and one from the bottom was taken using a microscope equipped with a camera. After that, ticks were stored in tubes at -80°C until used for DNA extraction. The length of the alloscutum (a) and the width of the scutum(b) were determined and the scutal index calculated from their ratio (a/b) [351]. In a second step the idiosoma length was determined once by looking at the tick from the top and once from the bottom (to see how reproducible this measuring is) (A) and the scutal width (B) as well as the coxal gap (C) were also determined. In this way, a second value for the scutal index (A/B) and coxal index (C/B) were calculated[350]. As described in (Figure 6), feeding time of individual ticks was estimated based on scutal and coxal indices using the formulas determined by Gray et al. [350].

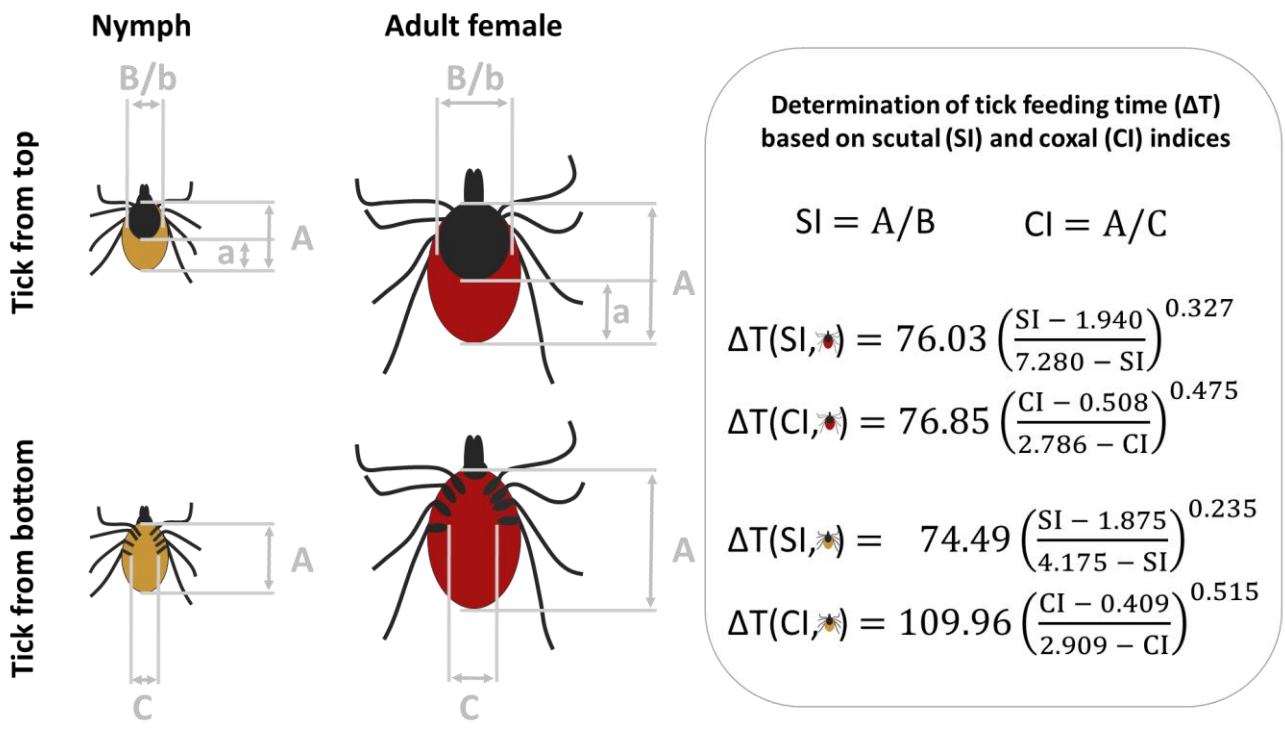


Figure 6: Schematic representation of method that was used to estimate tick feeding time. *A = idosoma length, B = scutal width, C = coxal gap, a = alloscutum length, b = scutal width. As indicated, a different formula was used for adult female ticks than for nymphs.*

3.3 Screening of ticks for the presence of *Borrelia* and *Rickettsia* pathogens

The PCR protocols used here have been setup by Anna L. Reye in the framework of a previous project from our laboratory[349]. Since ticks can harbour potentially dangerous pathogens, extraction of DNA was performed in a biosafety level 2 cabinet in our biosafety level 3 room. The QIAamp DNA Blood Mini kit (Qiagen) was used and the protocol of the manufacturer followed. Ticks were disrupted in 300ul AL buffer using the TissueLyser II (Qiagen). Amplification of *Borrelia* DNA (OspA gene) was performed in two consecutive steps using the primers from Michel et al. [329]. For the first round PCR, the following mastermix was prepared:

Reagent	**Volume/reaction [ul]**
ddH_2O	7.65
Buffer (10x)	2.5

$MgCl_2$ (50 mM)	1
dNTPs (10 mM)	0.5
SyBr Green (10X)	0.25
V1a primer (10 uM)	2
V1b primer (10 uM)	2
R1 primer (10 uM)	2
R37 primer (10 uM)	2
Platinum Taq® polymerase (5U/µl) (Invitrogen)	0.1
DNA extracted from tick	5
Total	25

To avoid contaminations, all the mastermixes prepared in this study were strictly pipetted in our clean mastermix room. DNA for first round PCRs was added in our PCR1 room in which only samples with low DNA concentration can be handled. For the first round OspA PCR, the following protocol was used:

Temperature [°C]	**Time [min:sec]**	**Cycles**
94	3:00	1
94	00:45	
50	00:45	40
72	01:00	
72	10:00	1

For the second round PCR the following mastermix was prepared:

Reagent	**Volume/reaction [ul]**

ddH_2O	11.65
Buffer (10x)	2.5
$MgCl_2$ (50 mM)	1
dNTPs (10 mM)	0.5
SyBr Green (10X)	0.25
V3a primer (10 uM)	2
V3b primer (10 uM)	2
R1 primer (10 uM)	2
R37 primer (10 uM)	2
Platinum Taq® polymerase (5U/µl) (Invitrogen)	0.1
Product from first round PCR	1
Total	25

DNA for second round PCR was added in our PCR2 room, in which more concentrated DNA samples and extracted plasmids can be handled. The following temperature protocol was used:

Temperature [°C]	**Time [min:sec]**	**Cycles**
94	3:00	1
94	00:45	
52	00:45	40
72	01:00	
72	10:00	1

Although this was not checked for in this case, this PCR should give a product of 850bps.

Detection of *Rickettsia* DNA (17 kDa) was also done by two rounds of PCR using the primers from Ishikura et al. [330]:

Reagent	Volume/reaction [ul]
ddH_2O	11.65
Buffer (10x)	2.5
$MgCl_2$ (50 mM)	1
dNTPs (10 mM)	0.5
SyBr Green (10X)	0.25
Rr17k.1p primer (10 uM)	2
Rr17k.539n primer(10 uM)	2
Platinum Taq® polymerase (5U/µl) (Invitrogen)	0.1
DNA extracted from tick	5
Total	25

The following temperature protocol was used:

Temperature [°C]	Time [min:sec]	Cycles
94	3:00	1
94	00:30	
55	00:30	40
72	00:45	
72	10:00	1

The following mastermix was prepared for the second round PCR:

Reagent	Volume/reaction [ul]
ddH_2O	15.65

Buffer (10x)	2.5
$MgCl_2$ (50 mM)	1
dNTPs (10 mM)	0.5
SyBr Green (10X)	0.25
Rr17k.90p primer (10uM)	2
Rr17k.539n primer (10uM)	2
Platinum Taq® polymerase (5U/µl) (Invitrogen)	0.1
Product from first round PCR	1
Total	25

The following temperature protocol was used:

Temperature [°C]	**Time [min:sec]**	**Cycles**
94	3:00	1
94	00:30	
54	00:30	40
72	00:45	
72	10:00	1

This PCR should yield a PCR product of 450bps. All PCRs on tick DNA from this study were performed in a 96 well CFX real time PCR machine (Biorad). Real time PCR curves were analyzed with CFX Manager (Biorad). For *Rickettsia* PCRs, products were analyzed on an agarose gel. Positive and negative controls were included in every PCR run.

3.4 Processing of patient blood samples and PBMC cryopreservation

Manipulation of patient material was done under biosafety level 2 conditions. PBMCs were extracted from 30ml of blood collected in sodium heparin tubes. We made sure to process the blood tubes within 8 hours after the blood draw. Tubes were stored at room temperature (RT) until processing. For peripheral blood mononuclear cell (PBMC) isolation, 15ml Ficoll was added to each 50ml Leucosep tube and centrifuged for 30-60s at 1000xg at RT. 15ml of blood were diluted with 15 ml of PBS and transferred to the prepared Leucosep tubes. After centrifugation for 40min at 400xg at RT without brake, the PBMC layer was removed and washed twice in 50ml PBS (300xg for 10min at RT). Cells resuspended in 5ml PBS were counted manually using Trypan blue staining and Neubauer chamber before another centrifugation step. Freezing was done by resuspending the PBMCs in 2.5ml cold complete RPMI1640 medium, followed by slowly adding droplet by droplet 2.5ml cold FBS containing 20% of DMSO [1]. After that, 3-5 aliquots of cells (1ml/tube) were prepared in precooled (-20°C) cryotubes and immediately put on ice and frozen down as quickly as possible in -80°C using a Mr. frosty. For long-term storage, cryotubes were transferred to a cryotank. Approximately 8ml of blood have been taken in serum tubes. After incubation for at least 30min at 37°C to allow blood clotting, the tubes were spun down for 30min at 3000rpm. The serum was aliquoted and stored at -80°C. Some of the blood samples used for PBMC extraction from seropositive and seronegative participants were collected in EDTA tubes.

3.5 Screening of patient sera for the presence of *Borrelia* antibodies

In general the last timepoint of each donor was tested for seropositivity against *Borrelia* using the commercially available kits for Europe from Sekisui Virotech. In case of a seropositive result, the first timepoint(s) was(were) also tested to check for a potential seroconversion. Besides this an in-house *Borrelia*-VlsE-C6 peptide ELISA has been developed using the following peptides as antigen: C(Mal-PEG2-Biotin)MKKDDQIAAAIALRGMAKDGKFAVK (VlsE-C6 epitope from *Borrelia bugdorferi s. s.* strain B31), C(Mal-PEG2-Biotin)MKKDDQIAAAMVLRGMAKDGQFALK (VlsE-C6 epitope from *Borrelia garinii* strain IP90) and C(Mal-PEG2-Biotin)MKKRNDNIAAAIVLRGVAKDGQFALK (VlsE-C6 epitope from *Borrelia afzelii* strain PT7) which have been synthesized by EMC microcollections with >90% purity. Peptide sequences

originate from [280]. VlsE peptides for ELISAs were stored at -80°C as 1-2mg/ml solutions in 50% DMSO. ELISA plates were coated overnight at 4°C with 50ul of 1ng/ml peptide solution in carbonate buffer and plates washed three times with washing buffer. Plates were blocked for two hours at RT with 150ul blocking buffer. After another three washing steps, 50 ul of diluted serum samples (1:100) were incubated for one and half hours before washing three times. The secondary antibody (Goat Anti-Human IgG-AP) was diluted 1:700 in dilution buffer and 50ul were incubated in the wells for one and half hours. The plates were washed an additional three times before substrate was added. Readout of the ELISA plate was performed at 405nm after incubation for one hour at 37°C. For ELISAs performed in 384 well plates a slightly changed protocol (with an easier way to prepare the buffers) was used. 2ug/ml in 20ul were used to coat the 384 well plates.

3.6 Multicolor flow cytometry

Cells were thawed as described in the next section. $2x10^6$ PBMCs were stained immediately after thawing in 100ul FACS buffer on ice. The samples were measured in three independent FACS experiments. One control sample was used for compensation controls and to test reproducibility between experiments. Compensation was done using a mix of cells and compensation beads. Antibodies were not titrated but the amount indicated on the vials was used. For multicolor staining, a mastermix containing all the antibodies was prepared. Control and acute samples were equally distributed among experiments. After addition of antibodies, cells were incubated on ice for 20-30min. 4ml FACS buffer was added to wash and the cells were centrifuged for 10min at 300xg and 4°C. Supernatant was removed by pouring and cells resuspended in 100ul FACS buffer. Hoechst was only added shortly prior to measuring on a FACSAria SORP (BD Biosciences) machine. Antibodies used in the panel: CD14-eFluor605NC, CD24-eF450, CD43-APC, CD23-APC-eFluor780 (eBioscience), IgD-BV421, CD19-BV605, IgG-PE (BD Pharmingen), CD10-BV510, CD138-BV711, CD27-PECF594 (BD Horizon), IgM-BV570, CD38-PerCP-Cy5.5, CD20-AF700, CD21-PE-Cy7 (BioLegend), CD5-FITC, CD3-PE-Dy647 (Immunotools). For the first experiment no CD43 staining was performed and the following antibodies were used instead in the corresponding channels: CD24-BV421 (BD Horizon), IgD-PacBlue and CD138-APC (BioLegend). Markers for this experiment have been selected with the help from Lynn Wenandy.

3.7 FACS staining for antigen-specific single cell sorting

Cells were thawed up on the day before the FACS sorting and isolated B cells were kept overnight in complete RPMI1640 medium at 37°C and 5% CO_2. For thawing, the cryovials were put into a water bath at 37°C until the outer layer of the ice crystal had melted. As soon as the ice crystal started melting, 1ml of prewarmed (37°C) complete RPMI1640 medium was added for complete thawing and cells were transferred into 8ml prewarmed (37°C) complete RPMI1640 medium. Cells were spun down at 300xg for 10min at room temperature, resuspended in 1ml medium and counted with Trypan blue staining and Neubauer chamber. The cells were kept in the incubator at 37°C and 5% CO_2 during the counting procedure. B cell isolation has been done by negative selection using the human B cell isolation kit II from Miltenyi. The separation of B cells was only suboptimal due to problems with cell clumping. This was probably a result of the long and too fast centrifugation after thawing.

The next morning, as described in [352] the biotinylated peptides and Neutravidin were mixed and incubated at a ratio 4 to 1 (each Neutravidin has four biotin binding sites). For each reaction, 1ul Neutravidin was mixed with 6.7ul of 10uM biotinylated peptide. The same peptides were used as for the ELISAs, but they were resuspended in water. The tube was incubated in the dark on ice for 20-30min followed by centrifugation for 10min at 4°C at maximum speed (to remove aggregates). Biotin was used as negative control as with the biotin binding sites blocked we saw less background than with untreated Neutravidin. In our hands, changing from Streptavidin [352] to Neutravidin reduced the background tremendously. During incubation, the isolated B cells from the previous day were spun down at 300xg for 10min and resuspended in cold 100ul FACS buffer. During the whole staining procedure, the cells were kept cold (either on ice or in the fridge). When the tetramers were ready, 7ul of tetramer mix were added to the cells and incubated for 30min on ice. For staining with mixed tetramers, the tetramers with each peptide were prepared separately and 7ul from each tetramer-peptide mix was added to each tube (meaning a total volume of 21ul was used also for the Biotin control tubes). After incubation, cells were washed twice (300xg, 10min) with 4ml of cold FACS buffer (keep FACS buffer all the time on ice or in the fridge). For first washing, 4ml FACS buffer were added to the 100ul of cells. For marker staining, an antibody mastermix was prepared and the appropriate amount (no titration was done, concentration indicated on antibody vials was used) was added to the washed cells after resuspending in 100ul FACS buffer. The following antibodies were used to distinguish the different memory B cell subpopulations and to gate out monocytes, T cells

and dead cells: CD14-FITC, CD3-FITC, CD20-Biotin (Immunotools), IgD-BV421, CD27-PECF594 (BDHorizon), CD19-BV605 (BDPharmingen) and Hoechst (Invitrogen). CD20-Biotin was used for compensation control for Neutavidin. After 30min incubation on ice, the cells were washed with 4ml of FACS buffer and resuspended in a minimum of 70ul (usually 100-200ul) FACS buffer for sorting on a FACSAria SORP (BD Biosciences) machine.

3.8 Single cell sorting

96 well PCR plates (Eppendorf) containing 5ul of 0,5x PBS, 10mM DTT and 5U Recombinant RNasin® Ribonuclease Inhibitor per well (according to single cell protocol from Buelow and Osborn [353]) were prepared in our mastermix room the day before the FACS experiment and frozen at -20°C for overnight storage. The plates were sealed with plastic foil. Before sorting the plates were taken out and put into the fridge for thawing. In order to make sure that the liquid is at the bottom of the wells and to avoid spilling due to removing the foil, a quick spin at 4°C has been done before putting them into the sorter. The plate holder of the sorter is kept at 4°C for the whole sorting procedure. Maximum speed at which cells were sorted was 3000events/second and maximum flow rate of 2 was used. Random or negative B cells were sorted into the first row of the plate and the rest of the wells were filled with peptide tetramer positive $CD19^{+}CD27^{+}(CD14^{-}, CD3^{-}, Hoechst^{-})$ B cells. In case of seropositive and seronegative donors, the sorting gates were set stricter, excluding also IgD^{+} B cells. The plates were sealed with plastic foil immediately after the sorting tube was empty and were quickly spun down (maximum 2min at 300xg and 4°C) to make sure that cells sticking to the sides of the wall of the wells also come into the liquid. In order to ensure RNA integrity, after sorting, the plates were immediately put on dry ice. A fresh plate was taken for every sample, no matter if the plate was full or not. After the experiment, the plates were transferred from dry ice to a -80°C freezer for storage.

3.9 Reverse transcription for single cell PCR

The PCR plate containing the stored single cells (5ul/well) was taken out of the -80°C freezer and always kept on ice when not in the PCR machine. Reverse transcription was performed in an

Eppendorf vapo.protect thermal cycler. In a first step, 9ul of cold reverse transcription buffer 1 was added to each well:

Reagent	**Volume/reaction [ul]**
First Strand buffer (2x)	5
Lysis buffer (5% Nonidet P-40/Igepal, Sigma)	3
Random primers (150ng/ul)	1
Total	9

The content of the plate was mixed by vortexing. A short spin of the plate was performed before putting it into the PCR machine at 65°C for 10min followed by 25°C for 3min. After these two temperature steps, the cells were put on ice for at least one minute before pipetting 5.5ul of the second mastermix containing:

Reagent	**Volume/reaction [ul]**
First Strand buffer (5x)	2
DTT (0.1M)	2
dNTPs (2.5mM)	1
SuperScript® III Reverse Transcriptase (200U/ul)	0.5
Total	5.5

After mixing and a short spin down, the plates were incubated at 37°C for 1h in the PCR machine followed by an inactivation step of the enzyme at 70°C for 15min.

3.10 Single cell PCR for amplification of immunoglobulin heavy chain genes

The single cell PCR was done using the primer sets from Wang and Stollar [331] but the restriction sites were omitted [332] from the primers. First round and second round PCRs were all performed in a total volume of 50ul in a UNO96 Thermal Cycler. For the heavy chain first round PCR the following mastermix was prepared in our mastermix room:

Reagent	**Volume/reaction [ul]**
PCR Buffer (10x)	5
dNTPs (2.5mM)	2
Primers (20pmol/ul) -> 0,4uM final	1(x9 different primers)
ddH_2O	25.75
HotStarTaq DNA Polymerase (Qiagen)	0.25
Single cell cDNA	8
Total	50

The single cell cDNA was added in our PCR1 room in which only low level DNA samples are allowed to be handeled. The following temperature protocol was used:

Temperature [°C]	**Time [min:sec]**	**Cycles**
95	15:00	1
94	00:45	
45	00:45	3
72	01:00	
94	00:45	
50	00:45	40
72	01:00	
72	10:00	1

For the heavy chain second round PCR, the same mastermix containing the second round primers and adjusted water was prepared and 3ul from the first round PCR was added in our PCR2 room in which more concentrated DNA samples are allowed to be handeled:

Reagent	**Volume/reaction [ul]**
PCR Buffer (10x)	5
dNTPs (2.5mM)	2
Primers (20pmol/ul) -> 0,4uM final	1(x9 different primers)
ddH_2O	30.75
HotStarTaq DNA Polymerase (Qiagen)	0.25
Product from first round PCR	3
Total	50

The following temperature protocol was used:

Temperature [°C]	**Time [min:sec]**	**Cycles**
95	15:00	1
94	00:45	
50	00:45	30
72	01:00	
72	10:00	1

No cells were sorted into the last row of the plates and these “empty” wells were used as negative controls for the PCR. In case the plate was not fully sorted, the remaining unsorted wells of the last sorted row were used as negative controls.

3.11 Single cell PCR for amplification of immunoglobulin light chain genes

For the light chain first round PCR the mastermix composition had to be adjusted:

Reagent	**Volume/reaction [ul]**
PCR Buffer (10x)	5
dNTPs (2.5mM)	4
Primers (20pmol/ul) -> 0,2uM final	0.5(x9 different primers)
$MgCl_2$	3
ddH_2O	25.25
HotStarTaq DNA Polymerase (Qiagen)	0.25
Single cell cDNA	8
Total	50

The same temperature conditions could be used as for the heavy chain first round PCR. For second round PCRs the kappa and lambda reactions were performed separately. The following mastermix was prepared for the amplification of kappa light chain sequences:

Reagent	**Volume/reaction [ul]**
PCR Buffer (10x)	5
dNTPs (2.5mM)	2
Primers (20pmol/ul) -> 0,4uM final	1(x5 different primers)
ddH_2O	34.75
HotStarTaq DNA Polymerase (Qiagen)	0.25
Product from first round PCR	3
Total	50

The same temperature protocol was used than for the heavy chain second round PCR. For amplification of lambda light chain products, the following mastermix was prepared this time using the primers from Tiller et al. [333,334]:

Reagent	**Volume/reaction [ul]**
PCR Buffer (10x)	5
dNTPs (2.5mM)	1
Forward primer mix (10uM)	1
Reverse primer (10uM)	1
ddH_2O	38.5
HotStarTaq DNA Polymerase (Qiagen)	0.5
Product from first round PCR	3
Total	50

The following temperature protocol was used:

Temperature [°C]	**Time [min:sec]**	**Cycles**
95	15:00	1
94	00:45	
57	00:45	40
72	01:00	
72	10:00	1

3.12 Analysis of correct sequence sizes on agarose gel and purification of PCR products for Sanger sequencing

In order to verify whether the PCR reactions were successful, 5ul of each PCR product were mixed with 1ul of 6x loading dye and loaded onto a 1.5% agarose gel. Depending on the number of samples, either a small or a large gel was prepared. 1.5g or 4.5g agarose were heated in 100ml or 300ml of TAE buffer and boiled in a microwave. From time to time the solution was mixed and boiling was stopped when the whole agarose was dissolved. The solution was cooled down to a temperature at which it was still liquid but which allowed to touch the bottle without any problems. To later visualize the DNA on the gel, 10ul or 30ul of SYBR®Safe DNA Gel Stain was added and the homogenized solution poured into the gel chamber. The solid and loaded gel was run at 130V for 30-45 minutes in a gel chamber containing TAE buffer. To visualize the DNA bands on the gel, it was illuminated in an inGenius gel documentation system and an image was saved with a GeneSnap image acquisition software. 1kb plus DNA ladder^{TM} was used as reference marker to estimate the size of the PCR products. The latter were purified using Jetquick PCR Product Purification Spin kit (Genomed Gmbh). In case of multiple bands the QIAquick Gel Extraction kit (QIAGEN) had to be used.

3.13 Sanger sequencing and analysis of single cell sequences

Sanger sequencing was performed in house by the technicians from our laboratory using the BigDye Terminator v3.1 Cycle Sequencing kit following the manufacturer's protocol on a 3130xl Genetic Analyzer. The same primers were used than the ones for the amplifying PCR. The raw sequences were analyzed and consensus sequences manually corrected by using SeqTrace [347]. IMGT/V-QUEST [341] was used to characterize the different sequences in terms of V and J gene usage as well as for CDR3 region identification. Clustering of single cell sequences was done using the bcRep R package 1.3.4 developed by Julia Bischof et al. [2] with the following settings clones95.tab<-clones(aaseqtab = aaseqtab, summarytab = summarytab, identity = 0.95, useJ = TRUE, dispCDR3aa = TRUE, dispFunctionality.ratio = TRUE, dispFunctionality.list = TRUE, dispSeqID = TRUE). In this example, the sequences having the same CDR3 length, V and J genes as well as a CDR3 identity threshold of 95% were grouped into clusters. This was repeated with several CDR3 identity thresholds ranging from 0% to 100%. Clustering of sequences based only on CDR3 sequence

similarity was performed using the Python script cl_1.1.9.py kindly provided by Jacob Galson [12,127] using the following settings python cl_1.1.9.py -l overview.txt -r $x -m 1 -t 1 -o output , where the following mismatches were allowed x=2 (50% CDR3 identity), x=3 (67% CDR3 identity), x=4 (75% CDR3 identity), x=5 (80% CDR3 identity), x=6 (83% CDR3 identity), x=12 (92% CDR3 identity).

3.14 RNA extraction from PBMC samples for library preparation

For thawing, best results (most viability of cells) was obtained when using the following procedure. Hold the cryovials into a water bath at 37°C until the outer layer of the ice crystal had melted. As soon as the ice crystal started melting, pour content of the vial into a 50ml falcon tube containing a mix of prewarmed (37°C) PBS and 3ml FBS. Centrifuge cells for only 5min at 1200rpm and resuspend them in 1ml PBS and remove aliquots for cell counting using Trypan blue staining and Neubauer chamber. Each tube contained frozen PBMCs from 6ml of peripheral blood. Spin the cells down for another 5min at 1200rpm during counting and immediately continue with DNA/RNA extraction using AllPrep DNA/RNA extraction kit (QIAGEN) according to manufacturer's description. Six extractions corresponding to the samples that were pooled on one Chip were performed at a time. RNA yield was quantified using a NanoDrop Spectrophotometer.

3.15 Library preparation and deep sequencing of the B cell repertoire

Next generation sequencing of patients' B cell repertoires was performed on an in house PGM Ion Torrent machine (Life technologies). The library preparation protocol of Vollmers et al. [1] using unique molecular identifiers (UIDs) in combination with BIOMED-2 V-gene FR2 primers [335] was adapted to the Ion Torrent system (Figure 7). Jean-Philippe Buerckert was so kind to share his mouse B cell repertoire protocol with me, which allowed to speed up the adaptation process. Barcoded primers were purified by PAGE while standard SePOP desalted primers were used for second strand synthesis. In the mastermix room, 1ul of 10mM dNTP solution and 2ul of 10uM barcoded isotype primers were added to a 96 well PCR plate, separating PCR reactions from each sample by at least one empty well in each direction. In the library preparation room, 500ng RNA for IgM and 1000ng for IgG was added to each well and the total volume adjusted to 9ul with RNAse/DNase free water.

In case of low RNA concentration, some reactions were done in two wells. The PCR plate was put into the PCR machine (UNO96 Thermal Cycler) programmed for 65°C 5min, 55°C 80min, 70°C 15min and a total volume of 20ul. When the PCR machine reached the 55°C step, 8ul of the following mastermix were added to each well without removing the plate from the PCR machine:

Reagent	**Volume/reaction [ul]**
First Strand buffer (5x)	4
DTT (0.1M)	2
RNaseOUT™ Recombinant RNase inhibitor (40U/ul)	1
SuperScript® III Reverse Transcriptase (200U7ul)	1
Total	8

The content was well mixed by pipetting up and down a few times. I designed the PCR plate layout in a way to have PCR reactions from the same sample (IgM and IgG samples from individual timepoints) in one row so that I could use the same capping strip for the same samples. This allows to minimize cross-sample contaminations to a maximum by always opening only wells together that contain the same sample. By preparing IgM and IgG samples in parallel, one has the option to control for how properly one has worked by checking how much IgM isotype sequences one can later find in the IgG samples and vice versa. After the reverse transcription, 30ul of the following second strand mix was added to each reverse transcription reaction:

Reagent	**Volume/reaction [ul]**
Phusion® HF Buffer (5x)	10
ddH_2O	7
V-region primer mix (10uM)	10
dNTPs (10mM)	1
DMSO	1.5
Phusion® High-Fidelity DNA Polymerase	0.5

Total	30

For second strand synthesis, tubes were put into the PCR machine that was programmed in the following way: 98°C 2min, 63°C 2min, 72°C 10min. After this, two consecutive manual Agencourt® Ampure® PCR purification bead (Beckman Coulter) cleanups were done using 1:1 bead to sample ratio and otherwise adhering to the manufacturer's instructions. In order not to lose any DNA, low DNA bind tubes (Eppendorf) were used for all the purification steps. The double stranded DNA was eluted in 21ul water from our water purifier of which 20ul were added to a new PCR plate. The following PCR reaction mix was subsequently added:

Reagent	**Volume/reaction [ul]**
Q5® Reaction Buffer (5x)	10
Q5® High GC Enhancer (5x)	10
ddH_2O	4.5
IonTorrent adapter primer mix (10uM)	2.5
dNTPs (10mM)	2
Q5® Hot Start High-Fidelity DNA Polymerase	1
Total	30

The PCR was performed in the same PCR machine (UNO96 Thermal Cycler) as used before, this time using the following temperature protocol:

Temperature [°C]	**Time [min:sec]**	**Cycles**
98	5:00	1
98	00:10	
65	00:20	19-20
72	00:30	

72	02:00	1

After this, the libraries were purified once using Agencourt® Ampure® PCR purification beads again at a bead to sample ratio of 1:1 before analyzing the quality and the quantity on a 2100 Bioanalyzer Instrument (Agilent Technologies). Those libraries that passed the quality check were pooled and a second Agencourt® Ampure® PCR purification bead cleanup of the pools was usually performed to completely get rid of potentially disturbing primer dimers before quantifying the pool on the 2100 Bioanalyzer Instrument. The pools were sequenced using Ion PGM™ Template OT2 400 Kit, Ion PGM™ Sequencing 400 Kit and ION 318 CHIP KIT V2 (Life technologies) according to the manufacturer's instructions. Because the old kits were no longer available at that time, the last sequencing run was performed using the newer Sequencing kit version (Ion PGM™ Hi-Q™ Sequencing Kit). After setting up the first deep sequencing runs for this project, Regina Sinner was so kind to introduce me into this technique so that I could perform the deep sequencing experiments myself.

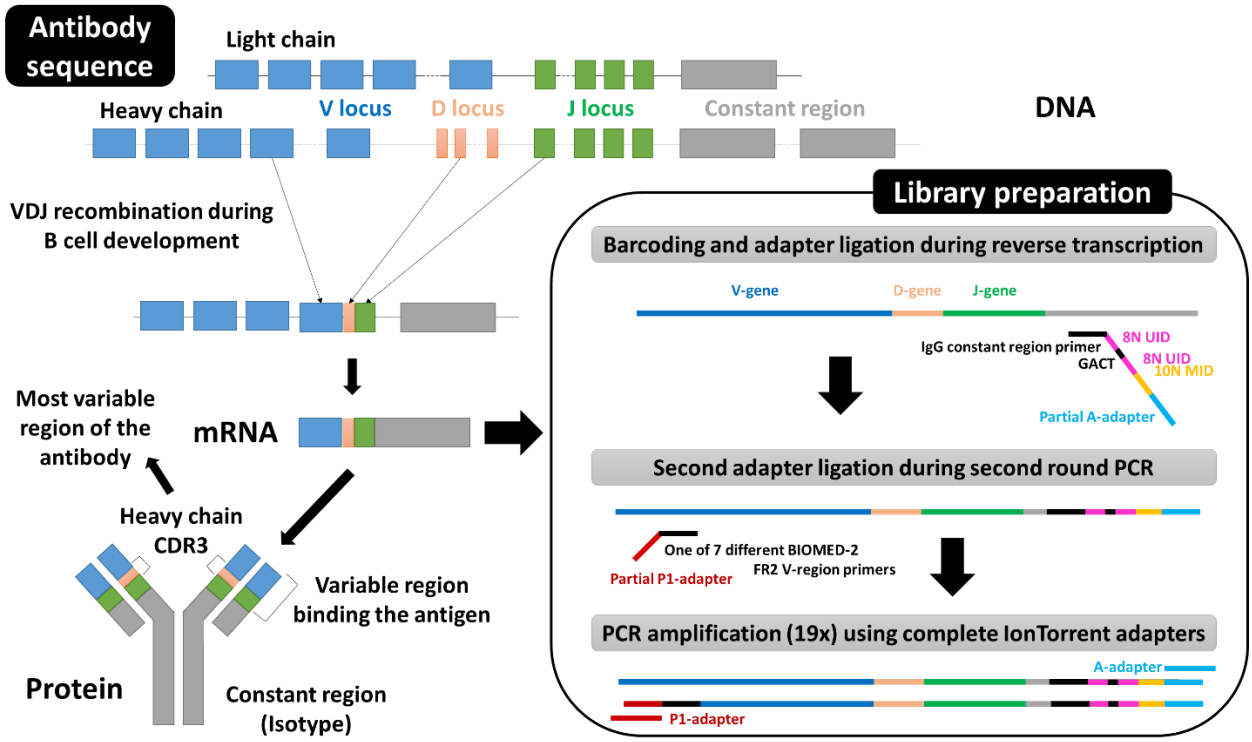


Figure 7: Schematic representation explaining library preparation method. *In our protocol, each mRNA is labeled with a unique molecular identifier (UID) barcode (highlighted in pink) before PCR amplification, allowing to assess PCR bias and coverage of the B cell repertoire. By generating consensus sequences from multiple reads, errors can be corrected as well. In addition this method*

only uses specific primers in the first amplification cycle and allows to reduce the primer sets to only one primer pair for the actual PCR amplification. This allows to reduce PCR amplification bias by avoiding multiplexing of different primers. As the Ion Torrent system is mainly doing insertion/deletion errors, we assume that selection of sequences containing exact matches of MID-NNNNNNN-GACT-NNNNNNNN-primer at the beginning of their sequence enables us to select for the majority of error free UID sequences from our samples and allows to discard reads having insertion/deletion errors at this site. Because the error rate increases with sequence length, the barcode was put at the beginning of the sequence. We start the sequencing from the constant part of the immunoglobulin heavy chain, as like this the CDR3 region is the closest possible to the beginning of the sequence. This enables us to get the maximum quality possible at the region we are the most interested in. In total 6 different molecular identifier (MID) barcodes (highlighted in orange) have been used to distinguish samples that could thus be pooled on the same Chip (IonXpress MID09: TGAGCGGAAC, MID15: TCTAGAGGTC, MID21: TCGCAATTAC, MID22: TTCGAGACGC, MID23: TGCCACGAAC and MID25: CCTGAGATAC).

3.16 Flow cytometry and deep sequencing data analysis

Kaluza Flow Cytometry Analysis Software was used for multicolor flow cytometry data analysis. Graphs and figures of this study were made using Power point and Graphpad Prism 5 Software. Mean and Standard Error of the Mean are represented in the different graphs. Significane: ****: $p < 0.0001$; ***: $p = 0.0001$-0.001, **: $p = 0.001$-0.01; *: $p = 0.01$-0.05. Pre-processing of deep sequencing rawdata was performed using an in house cleanup pipeline developed by labmates (please refer to papers and thesis of Jean-Philippe Buerckert and Axel R.S.X. Dubois from our laborartory for more details). William Faison was so kind to do the data cleanup for this project. Briefly, unaligned and untrimmed raw bam files were downloaded from the Ion Torrent server and converted to fastq and fasta formats. Only sequences with correct and good quality MID and UID regions were kept for further analysis. Those were selected by searching for the following pattern without allowing any mismatch or frameshift: MID sequence – UID1 sequence – GACT – UID2 sequence – primer. Only reads for which at least 80% of the bases contained a quality score above 20 were kept for further analysis. Any reads exceeding 420bps in length were cleaved off after this threshold and all reads smaller than 180bps were discarded. After that, consensus sequences were generated using a multithread python script in combination with IgBlast [339]. For each UID, sequences defined as unproductive by IgBlast were excluded from the alignments and a consensus sequence was subsequently built with only productive reads using PAGAN graph aligner [342]. Singlet sequences were discarded from the analysis. After that, IMGT/High V-QUEST [354] was used for selection of productive final consensus sequences, assignment of V, J and CDR3 sequences as well as determination of V-gene similarity to germline genes. VDJtools [348] was used for analysis of

different general repertoire parameters, including CDR3 length, V and J gene distributions as well as repertoire diversity. Spectratype distortion was calculated by determining standard deviation of CDR3 length distribution using popData in R as in [20]. Clustering of sequences was performed using Change-o [338] kindly provided to labmates by Steven Kleinsteins group and the Python script kindly provided by Jacob Galson (which was also used for single cell data analysis) [12,127,171]. William Faison performed clustering analysis using the Change-o pipeline for this project. Unless otherwise indicated, deep sequencing data analysis was performed on final uncollapsed productive consensus sequences. Only phylogenetic tree analysis were performed on collapsed data. Standard bash scripting commands in Biolinux were used to handle the large datasets and to extract the information of interest from output files of the analysis tools. Grep and agrep (in case mismatches were allowed) commands were used to check for the presence of sequences of interest in samples. Cd-hit [336,337] was used to cluster CDR3 sequences or UIDs according to different degrees of similarity.

3.17 Phylogenetic tree analysis of individual B cell clones of interest

After selecting all the sequences belonging to the clone of interest (extracted from Change-o output files), they were aligned manually using BioEdit and cleaved off at the correct position (please refer to Figure 42 in the results section). Remaining insertion and deletion errors were manually corrected (please refer to the second paragraph of the following discussion point: 5.2 Strengths and weaknesses of the approach used herein to identify *Borrelia* specific B cell repertoire signatures) before collapsing and counting nucleotide sequences. Phylogenetic trees were generated from the corresponding amino acid sequences (generated by IMGT-VQUEST [341]) using the one click approach from phylogeny.fr[343,344]. The germline sequences from phylogenetic trees in this study were determined manually by using the first V germline sequence determined by IMGT [341] until the V-D-J junction. As junction, generally the sequence present within the majority of sequences from that clone was taken. Whenever the D gene was clearly identifiable, this one was used, but this was nearly never the case. After the junction, the sequence was continued with the germline J gene. This means, that the trees are mainly highlighting mutations that occurred within the V-gene but mutations within the CDR3 regions were difficult to assign and are therefore not really included.

3.18 *In vitro* single cell stimulation

The *in vitro* stimulation protocol was kindly provided by Elisabetta Traggiai and John Lindner (Department of Mechanistic Immunology, Novartis Institutes for BioMedical Research, Basel, Switzerland). Alessia Colone took over this part of the project and provided final ELISA results on positive wells from which B cell receptors could then be analyzed. Briefly, she first negatively selected B cells from an aliquot of frozen PBMCs with the Easy Sep Human B cell enrichment cocktail kit (Stemcell Techonologies). After that, IgG expressing memory B cells were negatively selected using FACS sorting on a FACSAria Sorp machine (BD Bioscences). The following antibodies were used to gate on $CD3^{-}CD14^{-}CD16^{-}CD56^{-}CD27^{+}CD20^{+}IgD^{-}IgA^{-}IgM^{-}$.B cells: anti-CD3 PE, anti-CD14 PE (Immunotools), anti-CD16 PercPcy5.5 and anti-CD56 PecPcy5.5 (Biolegend), anti-CD20 Alexa Fuor 488 (Biolegend), anti-CD27 BV421, anti-IgD PE (BD Biosciences), anti-IgA and anti-IgM (Jackson Immuno Research). After the sorting, cells were resuspended in supplemented RMPI 1640 and seeded in 384 wells plates at 1 cell/well mixed with $2{,}5x10^{5}$/ml of irradiated CD40L EL-4-B5 cells at 50 Gy and incubated in a stimulation cocktail similar to [355] allowing the production of antibodies. Cells were incubated at 37°C and 5% CO_2. Supernatants were collected after 12-13 days for ELISA screening. At the same time cells were lysed in a homemade lysis buffer that has the same composition than the one from the Dynabeads Oligo (dT)25 kit (Life technologies) and stored at -80 ^{0}C. Besides the above described *Borrelia* VlsE-C6 peptide antigens, Tetanus toxoid (TT) (Serum Institute of India), Cytomegalovirus Grade 2 (CMV), Varicella Zoster grade 2 (VZ), Measles Grade 2 (MV), Rubella K1S (K1S), Mumps Grade 2, Epstein-Barr-Virus Viral Capsid Purified (EBV), *Toxoplasma gondii* (TG) antigens (Microbix Biosystems Inc.) were also tested for reactivity by ELISA: ELISAs were performed in 384 well plates as described above, using 2 μg/ml antigen concentration for all the antigens, except for EBV from which 30 ng/ml was sufficient. To isolate RNA only from wanted wells, lysis buffer from the Dynabeads Oligo (dT)25 kit (Life technologies) was heated up to 95°C and was added in volumes of 30-90ul (the volume is dependent on how cold the plate still was) to thaw up the content of the well of interest by pipetting up and down. This was repeated until everything from the well could be transferred into a low DNA bind tube (Eppendorf). The tube was filled up to 300ul of lysis buffer and RNA extracted according to the manufacturer's instructions. At the end Dynabeads to which the extracted mRNA is

attached were washed with 20ul 1x First Strand Buffer. They were kept in First Strand Buffer on ice until the following maxtermix was ready to be added:

Reagent	**Volume/reaction [ul]**
dNTPs (10mM)	1
5x First-strand buffer	4
rRNasin RNase inhibitor (40U/ul)	0.5
DTT (0.1M)	1
ddH_2O	13
SuperScript® III Reverse Transcriptase (200U/ul)	0.5
Total	20

The First Strand Buffer was removed from the beads before they were resuspended in this 20ul mixture and incubated for 45 min at 50°C in a heating block. On completion of the reverse transcription, the beads were washed with 20 µl of 1x First Strand Buffer. Since the cDNA is attached to the beads, thorough homogenization of the solution prior to PCR amplification is required. Immunoglobulin PCRs were carried out immediately after cDNA synthesis as described above, with the exception that for the first round PCR only 3ul input material were used instead of 8ul. Beads containing cDNA were kept in the fridge for storage.

4 RESULTS

4.1 Antibody responses in selected patients

A detailed description of the donors that have been selected for closer analysis of the B cell immune response to *Borrelia* can be found in Table 3 (acute donors), Table 4 (tick bite donors) and Table 5 (chronic patients). Seropositivity and seronegativity of the participants was determined in two ways. Besides the recommended two-tier method, for which we used commercial ELISAs and Immunoblots adapted for Europe, we also used an in house IgG VlsE-C6 peptide ELISA. In the latter, reactivity against peptides from three different *Borrelia* strains (B31, IP90, PT7) was assessed. We were able to confirm a strong correlation between reactivity of sera against the VlsE-mix on the commercial immunoblots and our in house ELISAs (Figure 8). Although the latter seems less sensitive, we were able to confirm, that all but one (95%) of the donors that strongly reacted with the immunoblot were also positive for at least one of the three peptides analyzed with the in house ELISA, confirming the completeness of the selected peptide pool (Figure 8).

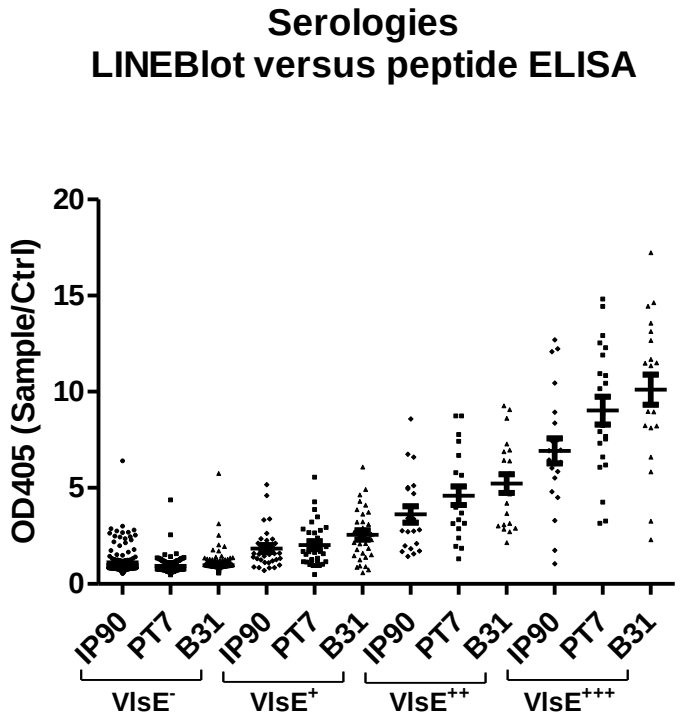


Figure 8: Correlation between VlsE reactivity in the commercial IgG immunoblot and in our in house VlsE-C6 IgG peptide ELISA. *Peptides from three Borrelia strains (IP90, PT7, B31) were tested. Samples were grouped according to their reactivity in the commercial immunoblot (negative to highly positive: VlsE^{-}, VlsE^{+}, VlsE^{++}, VlsE^{+++}) and the corresponding values obtained with the in house ELISAs plotted on the y axis. ELISA values were normalized to two negative control samples (Ctrl) measured on each plate. ELISA results were kindly provided by Alessia Colone.*

Interestingly, we saw a general trend towards a higher reactivity for the peptide originating from the american *Borrelia burgdorferi sensu stricto* strain B31 (Figure 9), a result that was confirmed in several experiments. Eight (50%) of the sixteen analyzed acute Lyme disease patients showed IgG reactivity with at least two of the three *Borrelia* VlsE-C6 peptides tested (Table 2). Five of the patients (31%) showed a change of reactivity over time for at least one of the three peptides (Table 2).

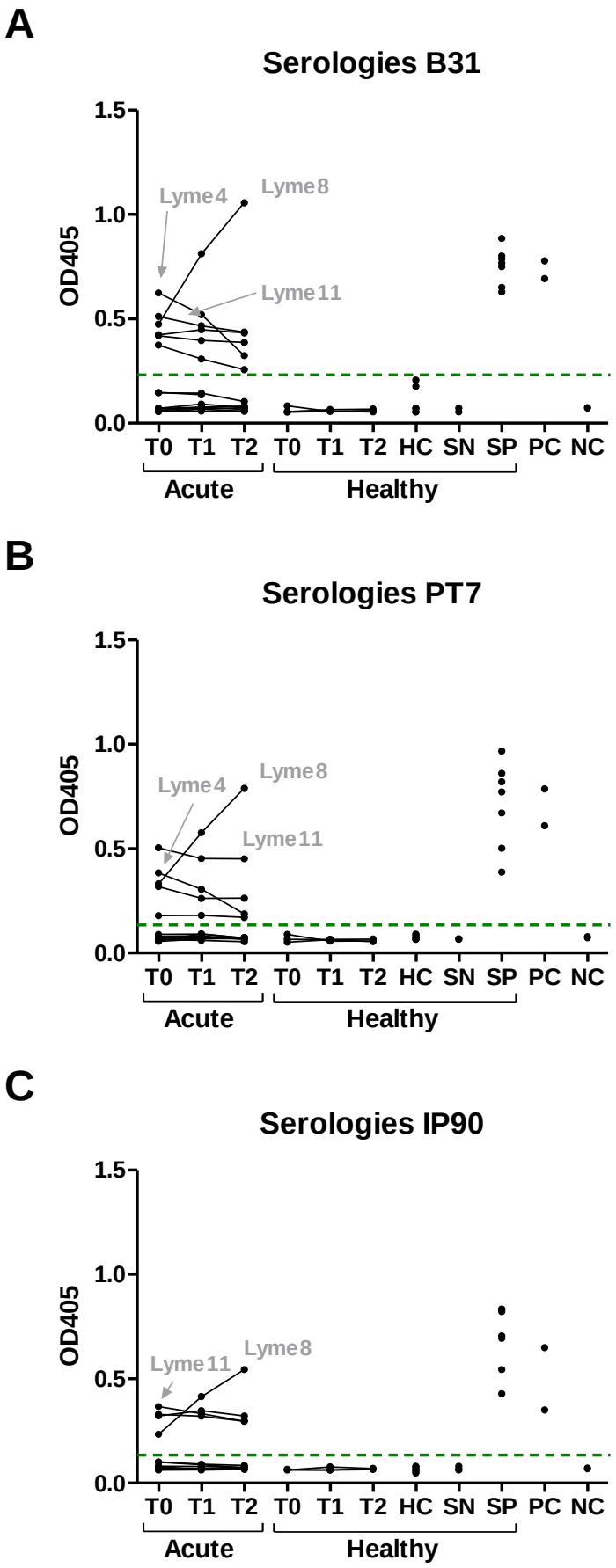

A
Serologies B31
OD405
1.5
1.0
0.5
0.0
Lyme4
Lyme8
Lyme11
T0 T1 T2 T0 T1 T2 HC SN SP PC NC
Acute
Healthy
B
Serologies PT7
OD405
1.5
1.0
0.5
0.0
Lyme4
Lyme8
Lyme11
T0 T1 T2 T0 T1 T2 HC SN SP PC NC
Acute
Healthy
C
Serologies IP90
OD405
1.5
1.0
0.5
0.0
Lyme11
Lyme8
T0 T1 T2 T0 T1 T2 HC SN SP PC NC
Acute
Healthy

Figure 9: One example of the in house VlsE-C6 peptide ELISA assay used to test the serostatus of acute patients and controls. *Two donors showed a striking increase (Lyme8) or decrease (Lyme4) in antibody titers during the observation period. T0-T2: different timepoints, HC: healthy controls, SN: seronegative donors, SP: seropositive donors, PC: positive controls, NC: negative controls.*

Table 2: Overview of VlsE-C6 peptide ELISA results of individual Lyme disease patients. *Green: increase in antibody titer over time, Red: decrease in antibody titer over time, Blue: fluctuations in ELISA results but not really seropositive, Grey: stably seropositive donors, White: seronegative samples, Black: not determined. Numbers in brackets: fold change of antibody titer as compared to T0.*

	***Borrelia* peptides and time point**								
	B31			**PT7**			**IP90**		
Patient	**T0**	**T1**	**T2**	**T0**	**T1**	**T2**	**T0**	**T1**	**T2**
Lyme1	Neg$^{(1.0)}$	Neg$^{(0.54)}$	Pos$^{(1.68)}$	Neg$^{(1.0)}$	Pos$^{(3.05)}$	Pos$^{(2.50)}$	Neg$^{(1.0)}$	Neg$^{(0.27)}$	Neg$^{(0.80)}$
Lyme2	Pos$^{(1.0)}$	Pos$^{(1.27)}$	Pos$^{(1.28)}$	Pos	Pos	Pos	Pos$^{(1.0)}$	Pos$^{(1.64)}$	Pos$^{(1.66)}$
Lyme3		Neg	Neg		Neg	Neg		Neg	Neg
Lyme4	Pos$^{(1.0)}$	Pos$^{(0.83)}$	Pos$^{(0.52)}$	Pos$^{(1.0)}$	Pos$^{(0.80)}$	Pos$^{(0.49)}$	Neg	Neg	Neg
Lyme5	Neg	Neg	Neg	Neg	Neg	Neg	Neg	Neg	Neg
Lyme6	Neg	Neg	Neg	Neg	Neg	Neg	Neg	Neg	Neg
Lyme7	Neg	Neg	Neg	Neg	Neg	Neg	Neg	Neg	Neg
Lyme8	Pos$^{(1.0)}$	Pos$^{(1.71)}$	Pos$^{(2.23)}$	Pos$^{(1.0)}$	Pos$^{(1.74)}$	Pos$^{(2.38)}$	Pos$^{(1.0)}$	Pos$^{(1.77)}$	Pos$^{(2.32)}$
Lyme9	Neg	Neg	Neg	Neg	Neg	Neg	Neg	Neg	Neg
Lyme10	Neg	Neg	Neg	Neg	Neg	Neg	Neg	Neg	Neg
Lyme11	Pos	Pos	Pos	Pos	Pos	Pos	Pos	Pos	Pos
Lyme12	Pos	Pos	Pos	Neg	Neg	Neg	Pos	Pos	Pos
Lyme13	Pos$^{(1.0)}$	Pos$^{(0.82)}$	Pos$^{(0.69)}$	Pos	Pos	Pos	Neg	Neg	Neg
Lyme14	Pos	Pos	Pos	Pos	Pos	Pos	Pos	Pos	Pos
Lyme15	Neg	Neg	Neg	Neg	Neg	Neg	Neg	Neg	Neg
Lyme16	Neg	Neg	Neg	Neg	Neg	Neg	Neg	Neg	Neg

In the two-tier testings, where we tested for both IgM and IgG reactivity against *Borrelia*, a seropositive result was obtained for 13 (81%) of the acute Lyme disease patients. In six (38%) of the acute patients, a change in antibody reactivity was observed, determined either by a change in intensity of single bands in the two-tier test done in house or by comparing our results to previous diagnostic tests done outside by different laboratories. Seronegativity of control donors were confirmed before sequencing antigen-specific B cells or the bulk B cell repertoire. For phenotypic analysis, healthy status was considered for both seropositive and seronegative otherwise healthy donors.

Table 3: Patients for which acute Lyme disease has been diagnosed by a medical doctor from Luxembourg and which have been selected for analysis of the B cell immune response. *The assays that have been performed on the individual patient samples are indicated on both sides of the table. *appearance or change in intensity of the band over the three timepoints tested; • erythema migrans disappeared; [0]appearance or change of symptoms during treatment; [§]IgM ELISA performed outside was slightly positive, no LineBlot has been done; [#]IgM ELISA done outside was negative; [$]Significant change over time.*

		Symptoms				Serology				Lyme disease history and treatment				
	Patient ID	Picture of erythema migrans	Erythema migrans/facial palsy	Additional symptoms	Duration of symptoms (weeks)	Two-tier testing (Europe, Genzyme Virotech)	In-house VlsE-C6 IgG ELISA (B31)	In-house VlsE-C6 IgG ELISA (PT7)	In-house VlsE-C6 IgG ELISA (IP90)	Last tick bite	previous Lyme disease	Start antibiotics treatment (days)	No symptoms anymore at T2	
In vitro stimulation	Lyme1	Available	Erythema migrans	yes	<2 months	IgM: ?[$]; IgG: VlsE, (DbpA)*	+/-[$]	+/-[$]	+/-[$]	?	no (outside blood test done 17 days before was negative in IgM but positive in IgG)	17	no[#]	B-cell repertoire (NGS)
	Lyme2	Available	Erythema migrans	yes	<2 months	IgM?[$]; IgG: VlsE, p39*, DbpA, p58, p83	+[$]	+	+[$]	2.5 months	no (outside blood test done before was positive in ELISA for IgM and positive for IgG, however in this Immunoblot test no reactivity was yet present for p39, p83 and p58)	4	no•	Sequencing antigen-labelled B-cells
➡	Lyme3	Available	Erythema migrans	yes	<2 months	IgM: VlsE, IgG: VlsE, (p83)	-	-	-	1 month	no (outside blood test done 10 days before was negative)	10	no[#]	
➡	Lyme4	N.A.	Erythema migrans	yes	3 weeks	IgM: OspC, VlsE; IgG: OspC, VlsE	+[$]	+[$]	-	1 year	no (outside blood test done 8 days before was already positive for IgM and IgG)	11	no•	
	Lyme5	N.A.	Erythema migrans	no	2-3 months	Negative	-	-	-	1 month	yes	4	yes	
➡	Lyme6	N.A.	Erythema migrans	no	2 weeks	Negative	-	-	-	2 weeks	no	2	yes	
	Lyme7	Available	Erythema migrans	yes	1 week	IgM: VlsE, p39; IgG: VlsE	-	-	-	?	no	2	yes	
➡	Lyme8	Available	Erythema migrans	no	4 days	IgM: (VlsE)*; IgG: VlsE	+[$]	+[$]	+[$]	3 weeks	no	1	no[#]	
	Lyme9	N.A.	Erythema migrans	yes	<3 months	IgM: OspC, VlsE; IgG: (OspC), VlsE, (p58),(p83)	-	-	-	5 months	no (outside blood test done 9 days before was negative in IgM but positive for IgG)	0	yes	
	Lyme10	N.A.	Erythema migrans	no	?	IgM: OspC*, (VlsE)*	-	-	-	9 days	no	?	?	
	Lyme11	N.A.	Erythema migrans	yes	1 year	IgM: (OspC); IgG: VlsE, DbpA, (p58)	+	+	+	1 year	yes (3 years ago)	0	no[#]	Multicolour flow cytometry
	Lyme12	N.A.	Erythema migrans	no	2 months	IgM: VlsE; IgG: (VlsE)	+	-	+	2 months	outside blood test 3 months before negative	0	yes	
	Lyme13	N.A.	Erythema migrans	yes	1 month	IgM: OspC; IgG: VlsE	+[$]	+	+	?	?	9	no?	
	Lyme14	N.A.	Facial palsy		?	IgG: (OspC), VlsE, p39, DbpA, p58	+	+	+	?	?	?	?	
	Lyme15	N.A.	Erythema migrans	yes	2 days	IgG: (VlsE)	-	-	-	?	no	1	yes	
	Lyme16	N.A.	Erythema migrans	no	0-4 days	Negative	-	-	-	4 days	no	0	yes	

Table 5: Patients for which chronic Lyme disease has been diagnosed and which were included into the multicolor flow cytometry analysis.

	Symptoms	Serology			
Patient ID	Remaining symptoms	Two-tier testing (Europe, Genzyme Virotech)	In-house VlsE-C6 IgG ELISA (B31)	In-house VlsE-C6 IgG ELISA (PT7)	In-house VlsE-C6 IgG ELISA (IP90)
Chron1	loss of conciousness, tiredness, pain (joints), depression, loss of memory	IgM: (VlsE)	-	-	-
Chron2	tiredness, muscle and joint pain, neurological problems	IgM: (VlsE)	-	-	-
Chron3	?	Negative (was IgG positive two years ago)	-	-	-
Chron4	restricted capacity, atrial fibrillation, slow motion, reduction in psychological and physical endurance	IgG: (OspC), VlsE, p39, DbpA, p58,p83	+	+	+
Chron5	pain (nerves and joints)	IgM: OspC, IgG: VlsE	-	-	-

Table 4: Donors which participated because of a recent tick bite. *B cell repertoire deep sequencing analysis was performed on Tick 1-3, while Tick 3-5 were included into the multicolor flow cytometry experiments. CF: computational feeding time; SI: scutal index; CI: coxal index.*

	Tick bite			Serology				Lyme disease history
Donor ID	Pictures of the tick	Feeding time (h) (CF(SI)/CF(CI))	Time after Tick bite (d)	Two-tier testing (Europe, Genzyme Virotech)	In-house VlsE-C6 IgG ELISA (B31)	In-house VlsE-C6 IgG ELISA (PT7)	In-house VlsE-C6 IgG ELISA (IP90)	previous Lyme disease
Tick1		32(2.31)/24(0.69)	2	negative	-	-	-	no
Tick2		N.D.(1.82)/N.D.(0.32)	3	negative	-	-	-	no
Tick3		30(2.23)/11(0.55)	0	negative	-	-	-	no
Tick4		12(1.96)/17(0.60)	1	positive (IgM: OspC)	-	-	-	positive serology (~2006)
Tick5		37(2.48)/23(0.68)	1	positive (IgM: OspC, IgG: VlsE and DbpA)	+	+	+	2004

Twenty-nine ticks from participants were screened [349] for *Borrelia* and *Rickettsia*, the two most prevalent pathogens previously detected [349] in ticks from Luxembourg. Four (14%) ticks turned out to be positive for *Borrelia* (Figure 10) and two (7%) for *Rickettsia* DNA (Figure 11), values that closely match with previous findings (which were 16.3% and 6.7% respectively).

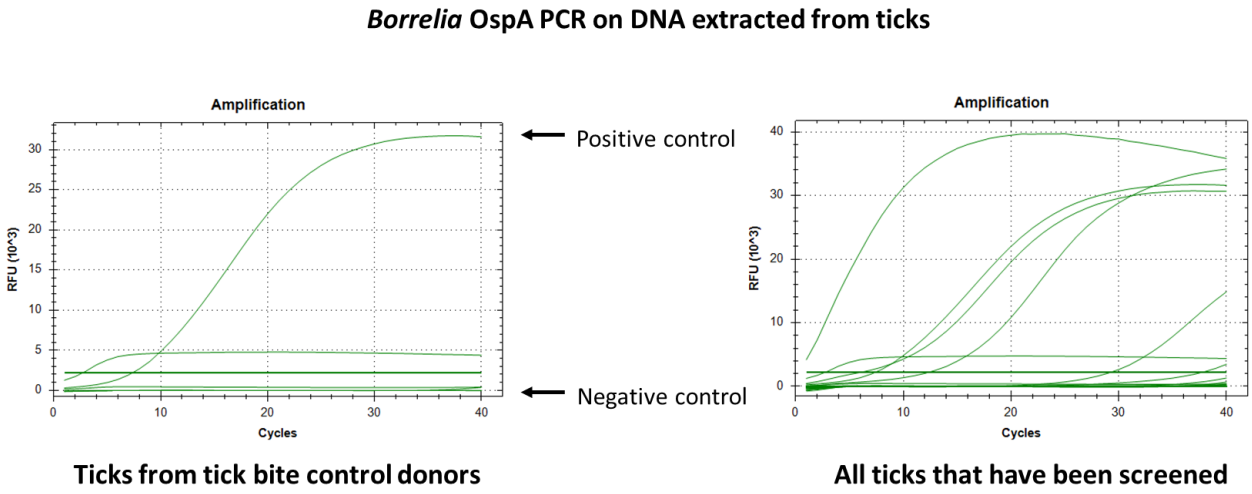


Figure 10: Results from OspA real-time PCR that was used to test for the presence of *Borrelia* DNA in ticks. *On the left, the results from the five donors used as controls in the deep sequencing and flow cytometry assays are depticted together with the positive and negative controls. On the right, the results obtained from all tested samples are depicted.*

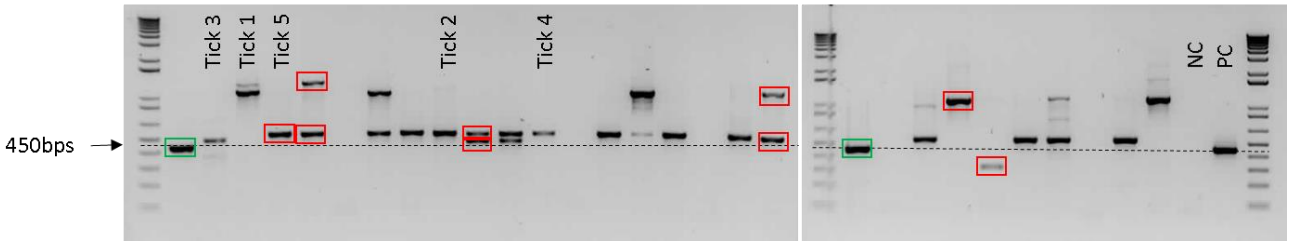


Figure 11: Results from 17-kDa PCR that was used to test for the presence of *Rickettsia* DNA in ticks. *Many side bands were obtained and the highlighted PCR products were sequenced. Only the green bands which also have the correct size can be considered as real positive for Rickettsia DNA.*

4.2 Analysis of patient's peripheral B cell subset composition with multicolor flow cytometry

To test whether acute Lyme disease affects peripheral B cell subsets, we developed a multicolor B cell panel (Figure 12). When analyzing the first six patients, we observed a rather constant donor specific peripheral B cell subset composition over time. No substantial changes, that would allow to distinguish acute Lyme disease condition from a healthy status could be determined from this low number of patients, indicating that occurring changes are rather of low level or very heterogeneous. In most cases, B cell subset composition was relatively stable throughout the one month observation period, confirming the reproducibility of our approach. After *Borrelia* infection, the time curve in particular of the plasmablasts (Figure 13E) and the activated B cells (Figure 13H) were much more

variable. The comparison of acute Lyme disease patients with controls revealed a significant increase in double-negative (Figure 13A) and IgG expressing memory B cells (Figure 13F). There was also a tendency towards higher levels of plasmablasts (Figure 13E) and activated naïve B cells expressing CD23 (Figure 13H) especially when comparing only to healthy controls. Interestingly in one of the donors with a recent tick bite, we observed a particularly skewed peripheral B cell subset composition. This donor had highly expanded non-switched memory B cells (Figure 13D) as well as $CD21^{low}CD38^{low}$ autoreactive B cells (Figure 13I) combined with relatively high levels of plasmablasts (Figure 13E).

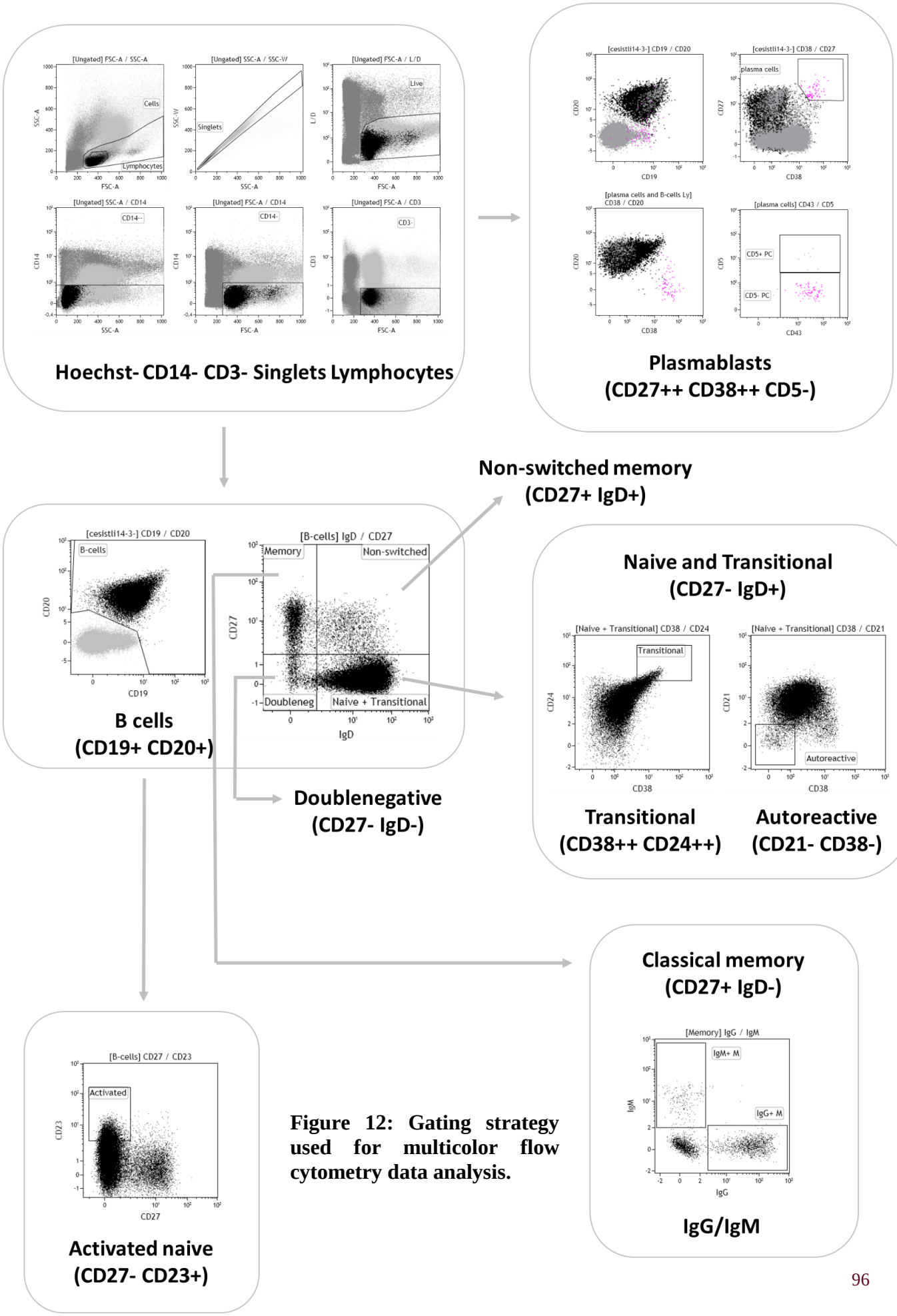


Figure 12: Gating strategy used for multicolor flow cytometry data analysis.

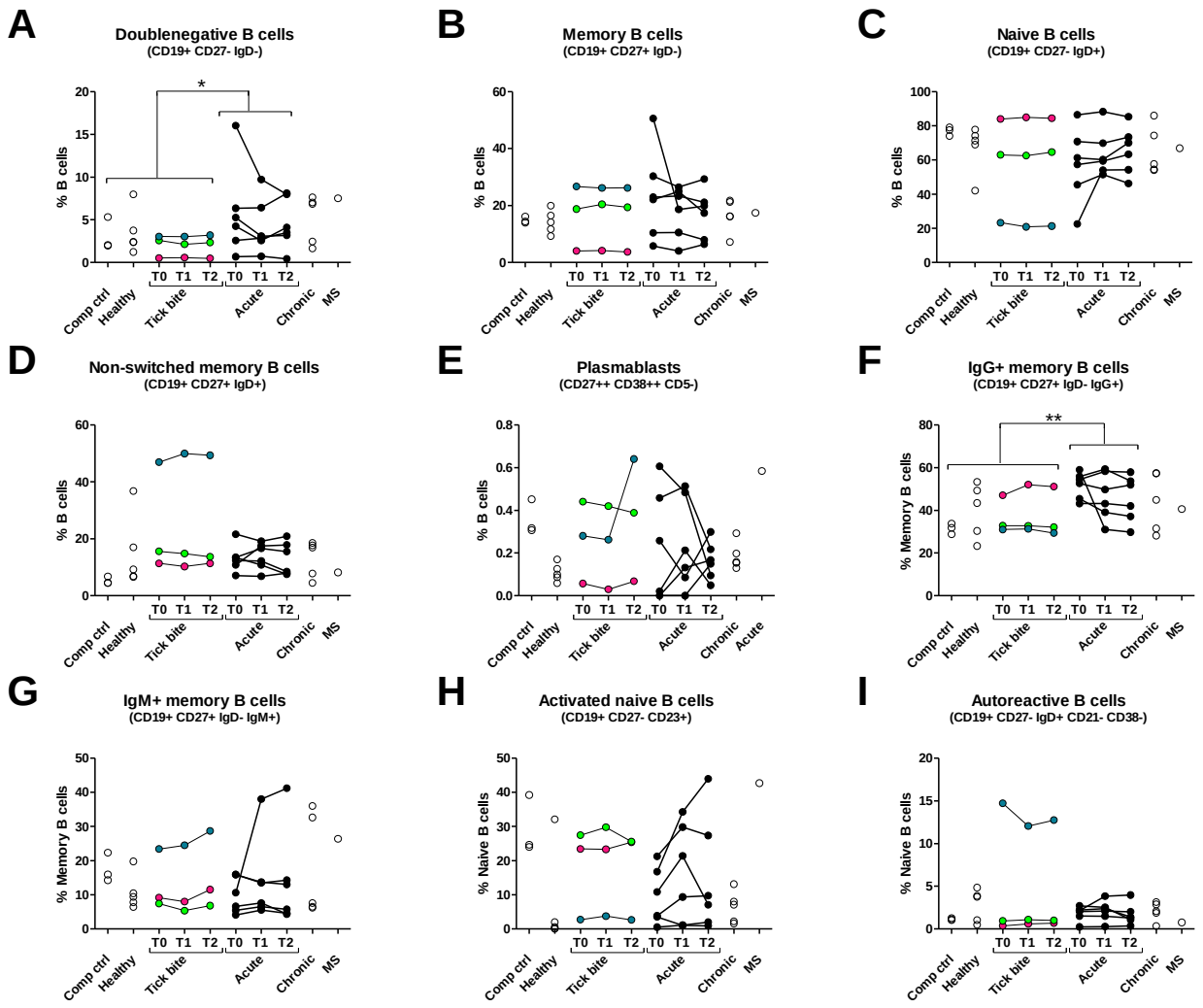


Figure 13: Results from multicolor flow cytometry experiment. *Comp ctrl: Sample used for compensation control and to assess reproducibility between the three experiments (each dot represents one experiment). Healthy: healthy individuals sampled at a single timepoint. Tick bite and Acute: Individuals with a recent tick bite and acute Lyme disease patients respectively sampled 3 times over one month. Chronic: patients for which chronic Lyme disease has been diagnosed. MS: patient for which at the initial visit Lyme disease has been diagnosed, but which was later found to suffer from multiple sclerosis. An unpaired two tailed t test with 95% confidence interval was used to determine statistically significant differences between samples of the two groups of main interest (control and acute samples combined).*

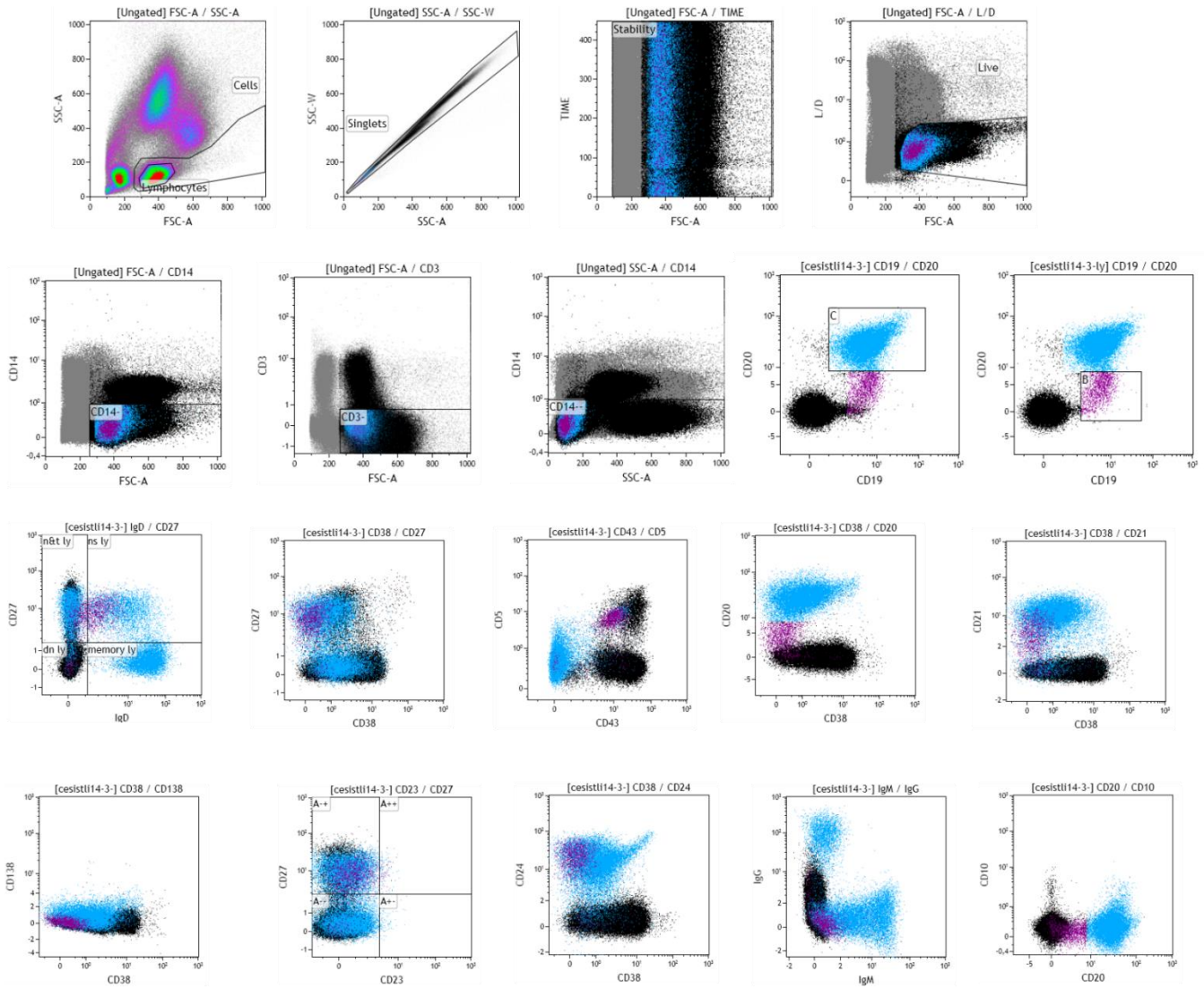


Figure 14: Description of a chronic lymphocytic leukemia (CLL) clone identified with multicolor flow cytometry. *In one acute Lyme disease patient (Lyme11) we detected a* $CD19^{dim}CD20^{dim}IgD^{dim}CD27^{pos}CD38^{neg}CD43^{pos}CD5^{pos}CD21^{dim}CD138^{neg}CD23^{neg}CD24^{pos}IgM^{neg}IgG^{neg}CD10^{neg}$ *subpopulation. After verification with the patient, it was confirmed that she suffered from a subclinical form of chronic lymphocytic leukemia. Blue:* $CD19^{pos}CD20^{pos}$ *B cells. Violet:* $CD19^{dim}CD20^{dim}$ *CLL clone. Black: other measured cells (PBMCs).*

In one of the acute Lyme disease patients, an unusual high number of $CD5^{+}$ B cells was observed. After closer analysis of this clone, it could be defined as a chronic lymphocytic leukemia (CLL) case. Another patient entered our study as acute Lyme disease patient, but in the course of the study it turned out to be rather a multiple sclerosis case. This patient is a good example for difficulties that can occur in the diagnosis of these complex diseases. For our multicolor flow cytometry panel it was a nice additional control. As peripheral B cell numbers and especially plasmablasts have been shown to be affected in this condition [356-360], we could confirm to find rather high numbers of these cells

in the peripheral blood of this patient (Figure 13E). Also other B cell subsets that are indicative of a rather active immune response were present at elevated levels in this donor. Activated naïve B cell (Figure 13H), but also doublenegative B cell (Figure 13A) and IgM expressing memory B cell (Figure 13G) percentages were relatively high in this donor.

To better understand changes in peripheral B cell subsets, we plotted the fold change of each subset compared to the general average on radar charts (Figure 15). This allows to represent the heterogeneous peripheral B cell subset compositions that we observed with our B cell panel. Half of the healthy individuals (Figure 15D-F) showed expansions of various B cell subsets (including activated naïve, transitional, doublenegative, non-switched memory B cells and plasmablasts) as compared to the average of all samples analyzed. The other three (Figure 15G-I) showed almost no levels of expansions of the different B cell subsets. The patient at a very early stage of the disease, similar to half of the control samples showed rather no B cell expansions (Figure 15J). Somewhat later, IgM^+ memory B cells and double-negative B cells were expanded (Figure 15K). In patients with an erythema migrans for 1-2 months, plasmablasts (Figure 15L) and IgG^+ memory B cell frequency (Figure 15M) was increased in combination with activation of naïve B cells. At later stages of the infection as facial palsy (Figure 15N) and erythema migrans since one year (Figure 15O) more differentiated subsets were not significantly expanded.

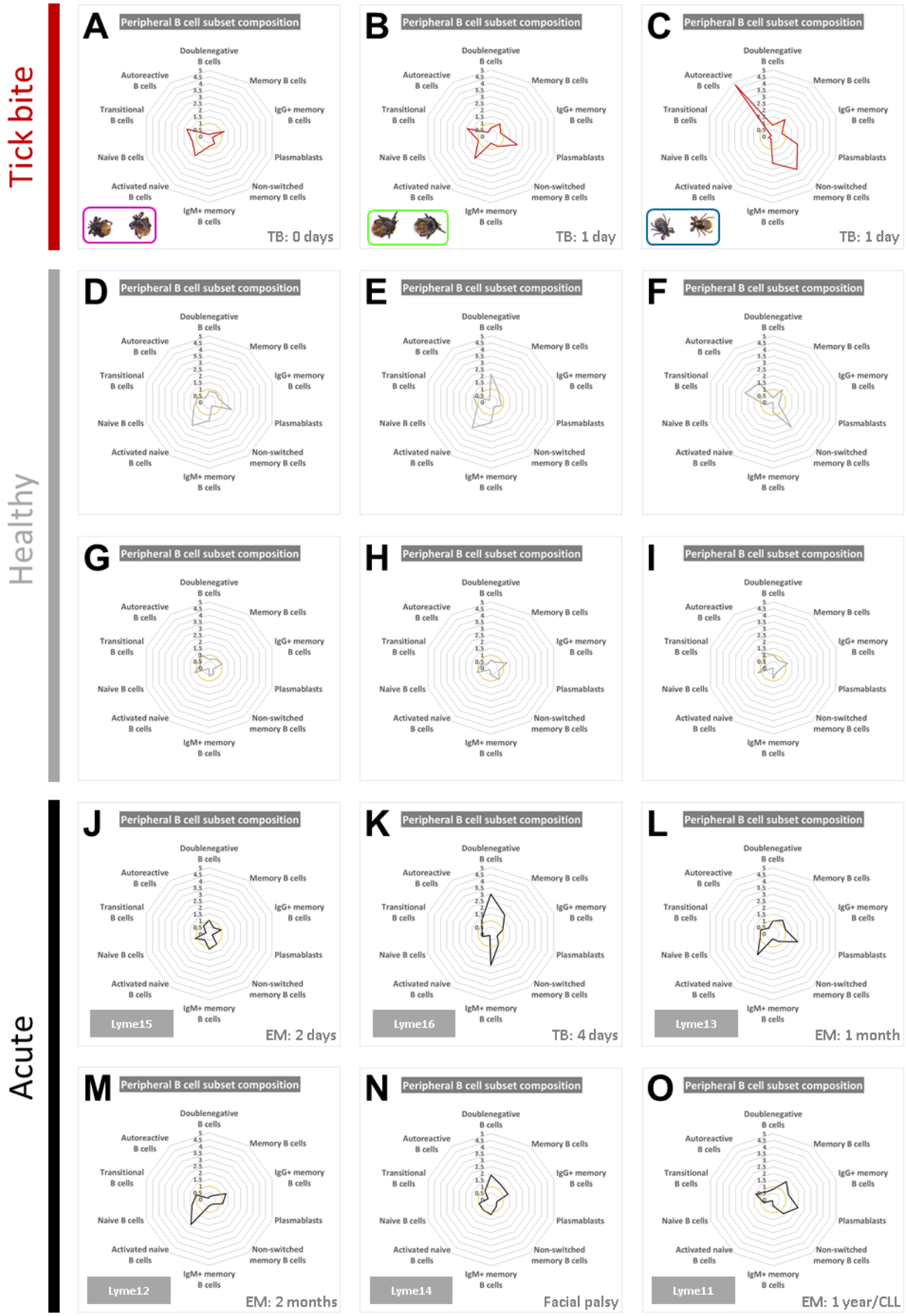

Tick bite
Healthy
Acute
A
Peripheral B cell subset composition
TB: 0 days
B
Peripheral B cell subset composition
TB: 1 day
C
Peripheral B cell subset composition
TB: 1 day
D
Peripheral B cell subset composition
E
Peripheral B cell subset composition
F
Peripheral B cell subset composition
G
Peripheral B cell subset composition
H
Peripheral B cell subset composition
I
Peripheral B cell subset composition
J
Peripheral B cell subset composition
Lyme15
EM: 2 days
K
Peripheral B cell subset composition
Lyme16
TB: 4 days
L
Peripheral B cell subset composition
Lyme13
EM: 1 month
M
Peripheral B cell subset composition
Lyme12
EM: 2 months
N
Peripheral B cell subset composition
Lyme14
Facial palsy
O
Peripheral B cell subset composition
Lyme11
EM: 1 year/CLL
Doublenegative B cells
Memory B cells
IgG+ memory B cells
Plasmablasts
Non-switched memory B cells
IgM+ memory B cells
Activated naive B cells
Naive B cells
Transitional B cells
Autoreactive B cells

Figure 15: Global analysis of changes in peripheral blood B cell subset composition using radar charts. *The fold change of each subset was normalized to the average of all the samples (highlighted as yellow line). Individuals with a recent tick bite are represented at the top (A-C). Healthy donors and acute Lyme disease patient samples are shown in the middle (D-I) and at the bottom (J-O) respectively. For both acute patients and donors with a recent tick bite average values of the three timepoints were calculated.*

4.3 Isolation of VlsE-C6 peptide reactive B cells with tetramer staining

Borrelia-reactive memory B cells were directly labeled and sorted using a peptide tetramer staining approach [352] with the immunodominant VlsE-C6 epitopes derived from the three most important *Borrelia* species [280]. To setup the staining, we started generating peptide tetramers using Streptavidin [352], but changed to Neutravidin which allowed us to solve problems with an initial high background noise. The peptide staining approach was validated by confirming an increased staining of switched memory B cells in seropositive as compared to seronegative donors in the absence of staining of naïve B cells (Figure 16). Also in acute Lyme disease patients, the memory B cells showed an increased reactivity (Figure 17). The staining pattern was similar to the one of seropositive donors, with only a subset of samples showing a detectable reactivity towards the peptide tetramers. We did not observe any correlation between reactivity towards the individual *Borrelia* peptides at the antibody and memory B cell levels (Figure 18).

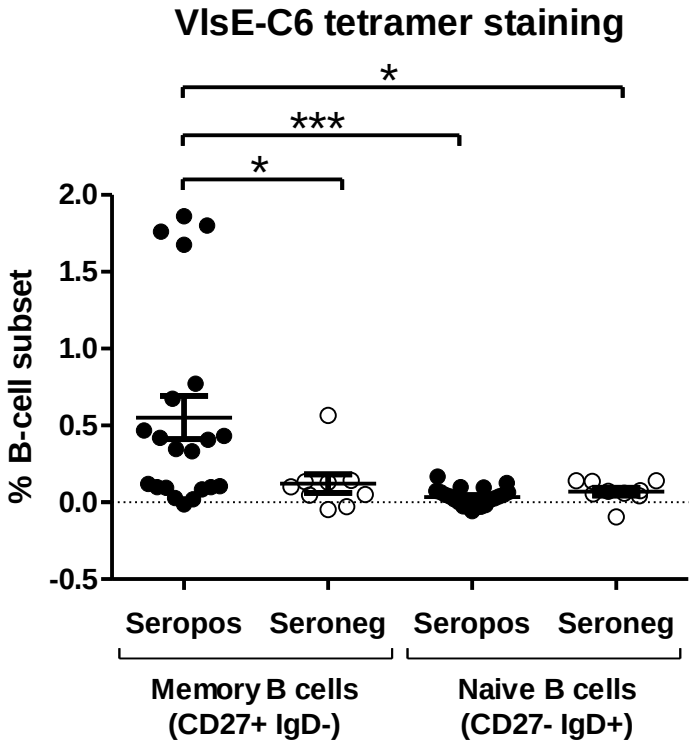


Figure 16: Labelling of memory ($CD27^+IgD^-$) and naïve ($CD27^-IgD^+$) B cells from seropositive and seronegative donors with VlsE-C6-Neutravidin tetramers to confirm staining specificity. *Background staining observed with biotin tetramers was subtracted to obtain the represented values. Reactivity of the samples from the different groups was compared with One-way Analysis of Variance test followed by Tukey's Multiple Comparison test. Represented are the results from five independent experiments including data from nine seropositive and five seronegative donors. Donor's samples were subdivided onto four (including biotin control) tubes and stained separately with tetramers containing peptides from three different Borrelia strains (B31, IP90, PT7). We don't distinguish values obtained with the different peptides in this figure as the same trend towards an increased staining of memory B cells in seropositive donors was observed with all of them.*

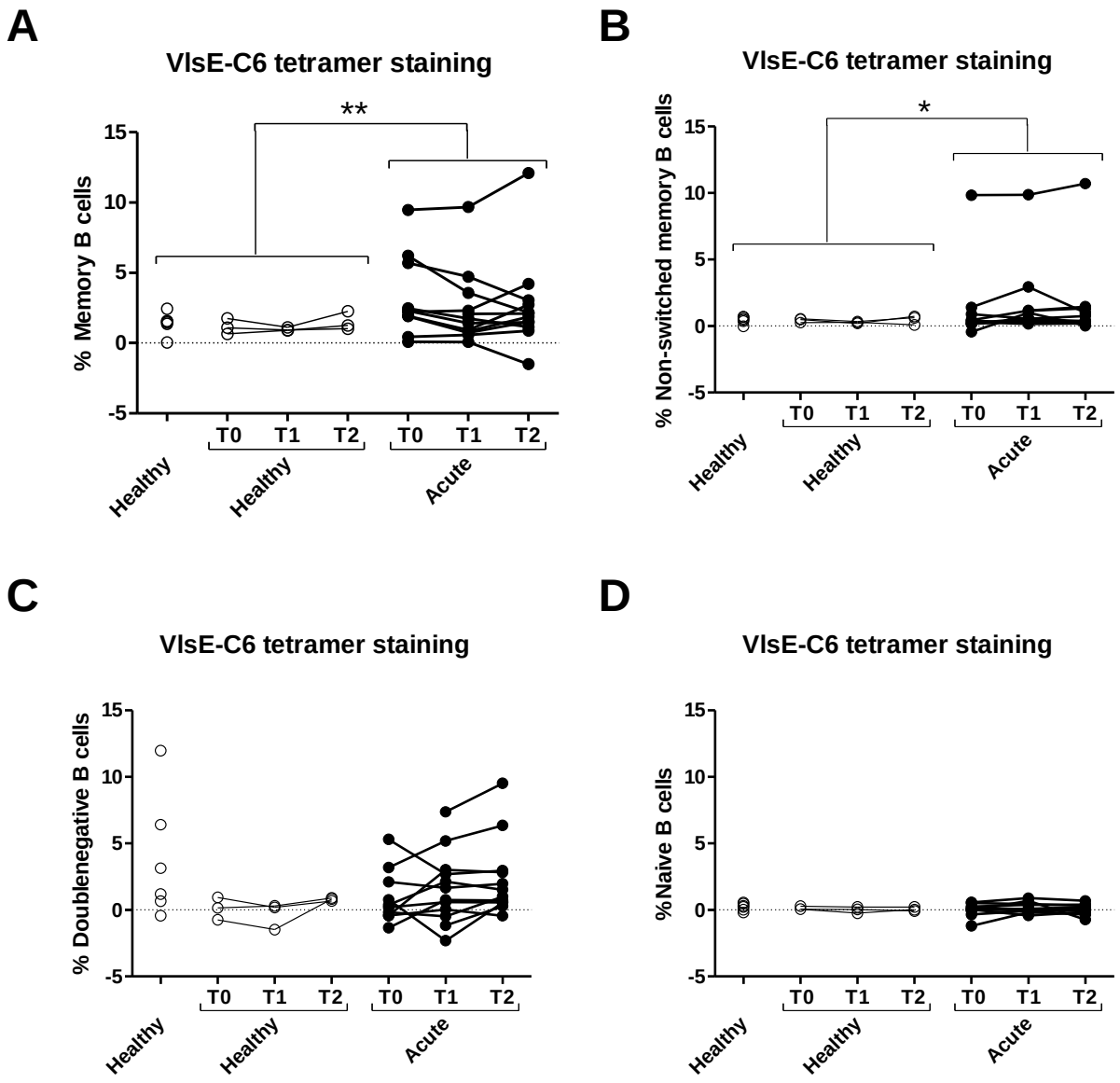
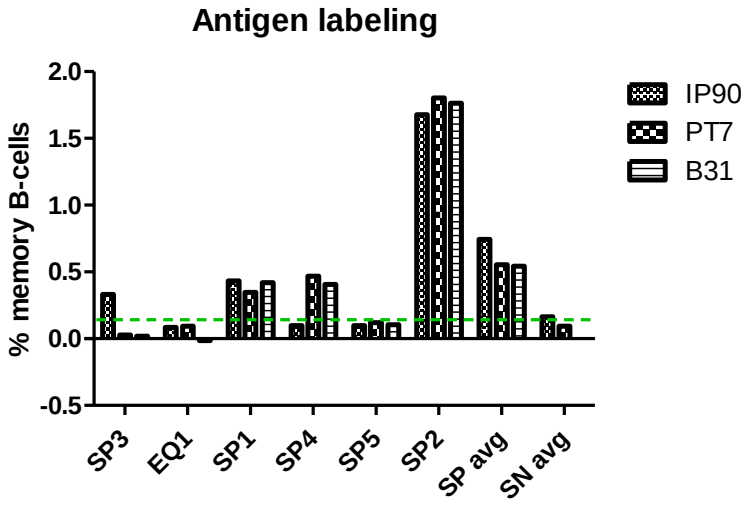
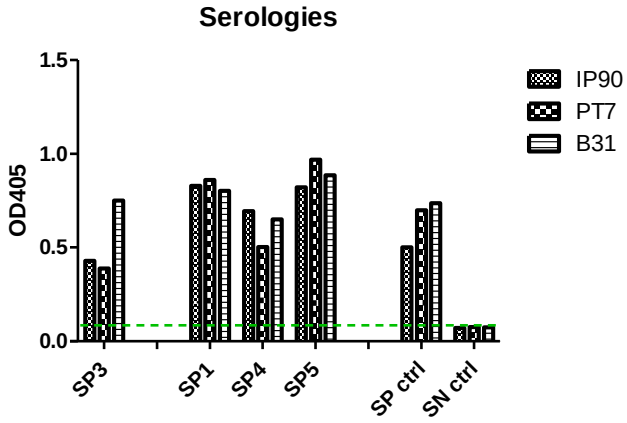


Figure 17: Labelling of VlsE-C6 peptide reactive B cells in acute Lyme disease patients. *Tetramers containing peptides from the three Borrelia strains (B31, PT7, IP90) were pooled and antigen reactive B cell numbers otherwise determined as in Figure 16. Only part of the healthy controls and all of the acute Lyme disease patients were sampled at different timepoints over one month period indicated as T0, T1 and T2. Samples from the two groups were compared using an Unpaired t test with Welch's correction. The observed difference was still significant when removing the outlying CLL patient for A (*) but no longer for B (n.s.). As described below, this high level could be explained by non-specific binding to the CLL clone.*

A


Antigen labeling
% memory B-cells
2.0
1.5
1.0
0.5
0.0
-0.5
SP3
EQ1
SP1
SP4
SP5
SP2
SP avg
SN avg
IP90
PT7
B31

B


Serologies
OD405
1.5
1.0
0.5
0.0
SP3
SP1
SP4
SP5
SP ctrl
SN ctrl
IP90
PT7
B31

C

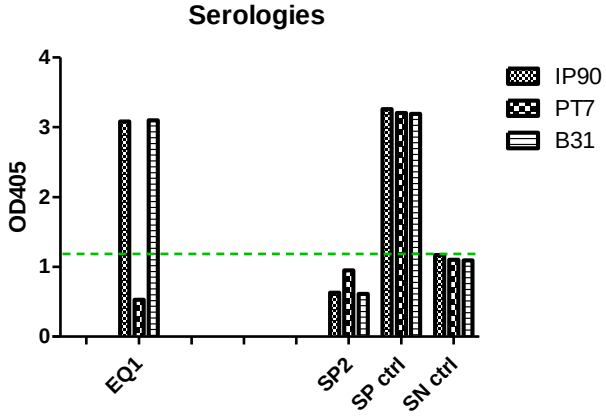

Serologies
OD405
4
3
2
1
0
EQ1
SP2
SP ctrl
SN ctrl
IP90
PT7
B31

Figure 18: Reactivity of seropositive donors towards the VlsE-C6 epitopes originating from three different *Borrelia* species at the memory B cell (A) and antibody (B-C) levels. *Due to differences in overall OD values, the results from the 96 well ELISA (B) assay are represented separately from those of the 384 well assay (B). The same control samples were used in both ELISA assays. The donor SP2 was only slightly reactive against the VlsE-mix on the more sensitive commercial immunoblot. This donor reported to have had a tick bite with a red skin lesion 5 weeks before blood draw. EQ1 was seronegative for VlsE-mix on the commercial immunoblot, but turned out to be seopositive by our in house VlsE-C6 peptide ELISA. B31: Borrelia burgdorferi sensu stricto; PT7: Borrelia afzelii; IP90: Borrelia garinii. SP avg: average of values from seropositive donors. SN: average values from seronegative donors. SP ctrl: average value from two seropositive donors. SN: average value from two seronegative donors. The green lines in each graph highlight the cut-off values, which were manually set.*

In one of the ten analyzed acute Lyme disease patients (Lyme11), we found a particularly large antigen-labeled B cell subpopulation. When analyzing the sequences from these cells, we found several different predominating expanded B cell clones. In addition to that, we found that these cells colocalized (meaning, they were $CD27^{+}IgD^{dim}CD20^{low}$) with the $CD5^{+}$ CLL B cell subpopulation that we previously detected in the same donor with our multicolor B cell panel (Figure 19). In order to test whether the CLL cells are indeed peptide reactive and to compare reactivity of the isolated B cell clones with the individual *Borrelia* epitopes, we labeled PBMCs separately with the three peptide tetramers in combination with the CD5 marker. As opposed to previous findings on MACS separated B cells for which around 10% of the memory B cells from Lyme11 were peptide positive (Figure 17), we were not able to reproduce this finding when using untouched whole PBMCs. Under these conditions, the CD5+ B cells were also not positive for peptide staining. Although labeling in this experiment was overall very low, we still saw a higher staining of memory B cells as compared to the naïve subset (Figure 20), confirming that a few peptide positive sequences should still be among the previously sorted B cells. When comparing the percentage labeled memory B cells we obtained when using the different epitopes, we found that the memory B cells of this patient most strongly reacted with the epitopes originating from *Borrelia burgdorferi sensu stricto* (B31) and *Borrelia afzelii* (PT7) (for memory B cells IP90: 0.03%, PT7: 0.09%, B31: 0.09% and for non-switched memory B cells: IP90: 0.02%, PT7: 0.05%, B31:0.07% staining difference as compared to naïve B cells) (Figure 20). At the serum level, this donor showed reactivity towards all three peptides (Figure 9). Interestingly, (even when leaving out the outlying Lyme11 sample) we observed a correlation between time of onset of symptoms and the number of tetramer reactive memory B cells (Figure 21).

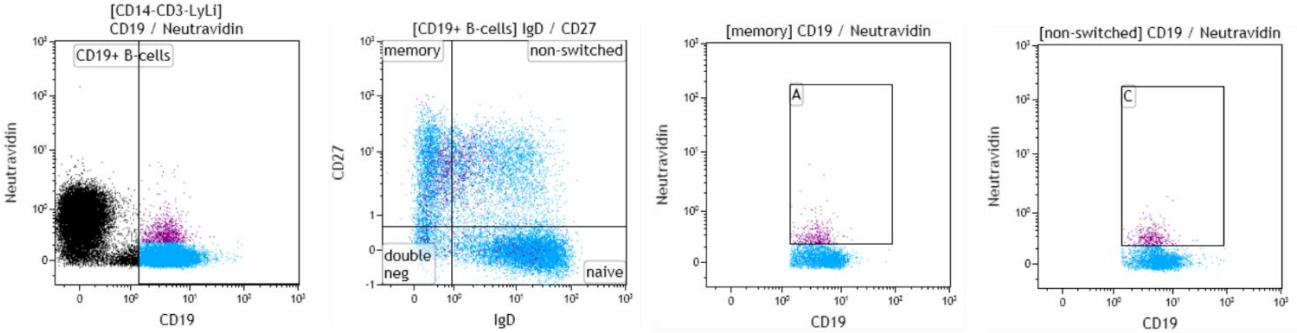


Figure 19: Peptide positive B cells in outlying donor (Figure 17, Lyme11). *The large subpopulation of peptide positive memory B cells measured in Lyme11 could be found at the same position (being $CD27^{dim}IgD^{dim}$) as the CLL clone previously characterized in the same donor (Figure 14).*

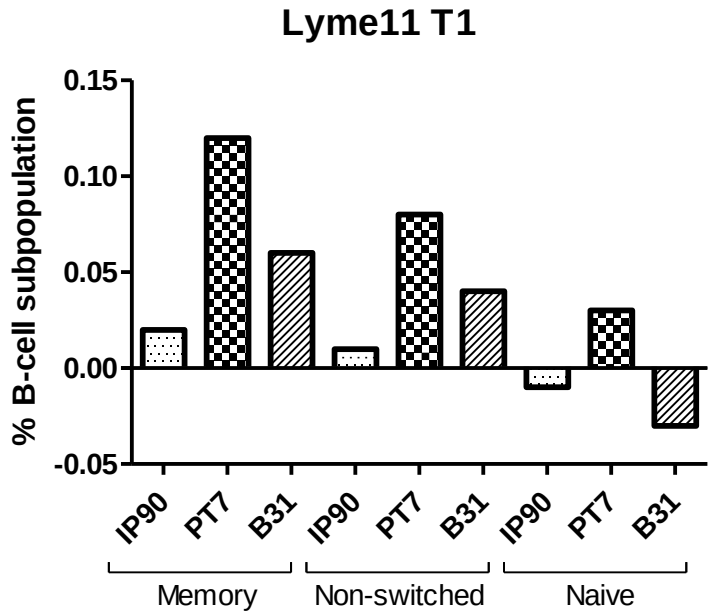


Figure 20: Repetition of tetramer staining with individual peptides on a sample from Lyme11. *This donor was highly positive for VlsE-C6 peptides in the first tetramer staining experiment. Staining was performed in the same way as in Figure 16, only that no MACS separation was done before labeling.*

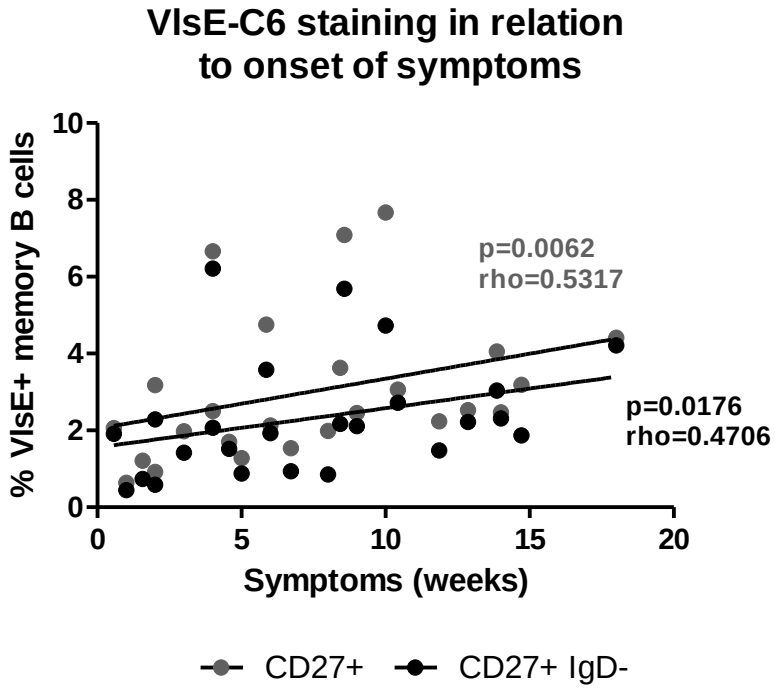


Figure 21: VlsE-C6- tetramer staining of total memory B cells (CD27+) and switched memory B cells (CD27+IgD-) versus the time that has elapsed from onset of symptoms are represented. *Spearman's Rank Correlation Hypothesis Testing was done manually in Excel.*

4.4 Single cell immunoglobulin heavy chain sequence cluster analysis

In order to characterize the antibody receptors from peptide positive B cells, single cells from 10 acute Lyme disease patients were sorted into 96 well plates and their heavy chains sequenced [331,333,334]. In total, we were able to get 628 sequences from peptide positive B cells and 362 sequences from control cells (including randomly sorted B cells, peptide negative and Neutravidin positive B cells) (Figure 22, Table 6). To compare the peptide reactive heavy chain sequences obtained from Lyme11 after MACS separation with the CLL cells, we also sorted and sequenced 52 single $CD5^+$ B cells from this donor (Table 6). When comparing these sequences, we found that they are the same clones and concluded, that these clones have nothing directly to do with Lyme disease, but have probably been sorted due to non-specific binding of the tetramers to $CD43^+$ CLL cells that have not been completely removed using MACS. Overall in this experiment MACS separation was not efficient, because of a large amount of cell death and clumping after thawing and centrifugations. We assume, that not all four biotin binding positions were occupied in our tetramers and that they therefore boud to remaining CLL cells attached to the anti-CD43-biotin antibody from the used

MACS kit. This would also explain the large amount of staining of cells other than $CD19^+$ B cells (Figure 22).

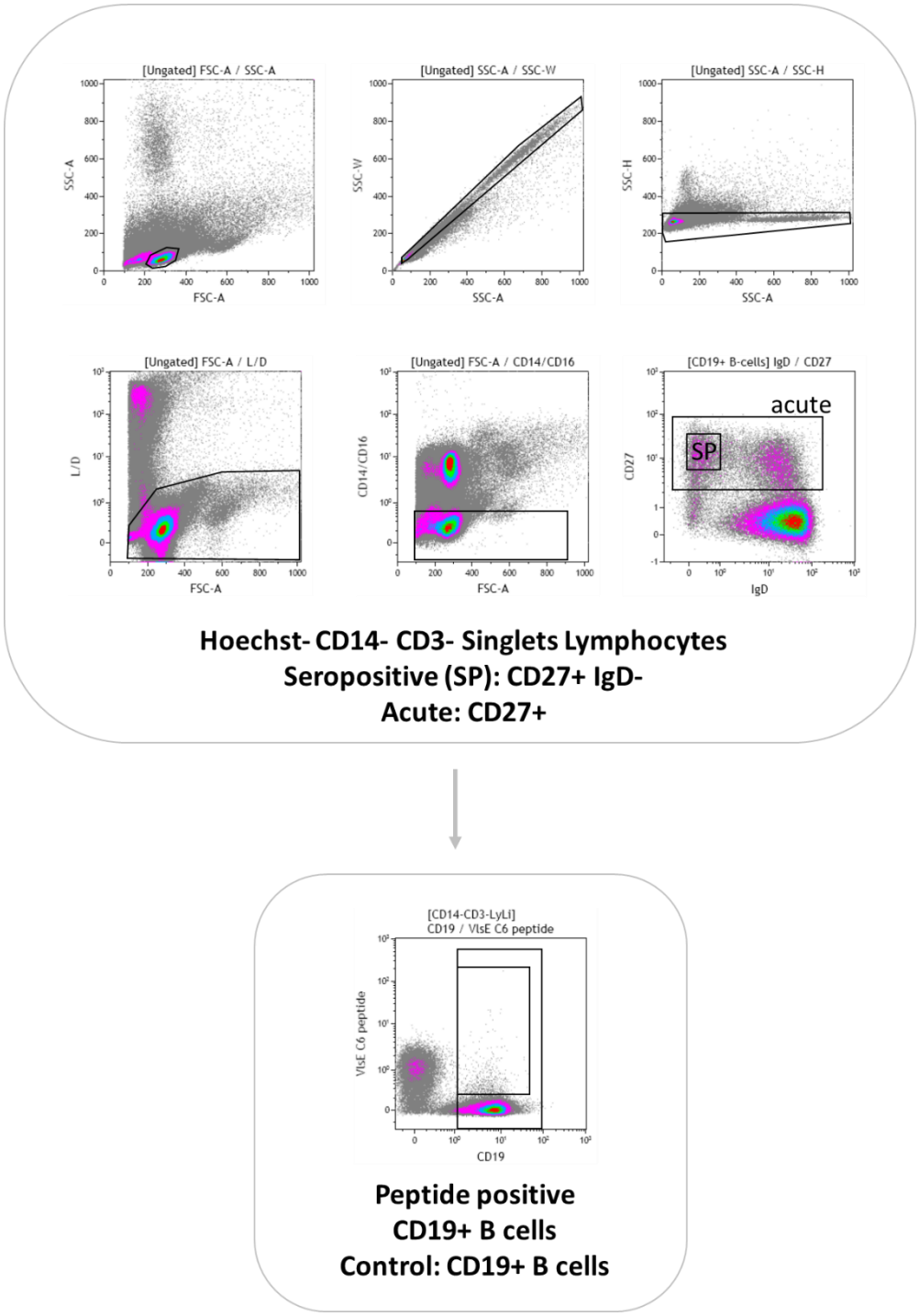


Figure 22: Gatings for measurement and sorting of VlsE-C6 peptide reactive memory B cells from acute and seropositive donors.

Table 6: Overview of number of single cells sequenced per donor.

Patient ID	Status	Random	Negative B31	Negative IP90	Negative PT7	Neutravidin	Peptide mix	B31	IP90	PT7	CLL
SN1	seronegative	8					19				
EQ1	equivocal		37			30		12			
SP1	seropositive		7	4	8			4	4	5	
SP2	seropositive		22	26	30			18	15	20	
Lyme3	acute LB	19					107				
Lyme4	acute LB	19					81				
Lyme5	acute LB	20					11				
Lyme6	acute LB	27					101				
Lyme7	acute LB	28					24				
Lyme8	acute LB	9					2				
Lyme9	acute LB	24					74				
Lyme10	acute LB	19					19				
Lyme11	acute LB	17					91				52
Lyme12	acute LB	8					21				

In order to distinguish between truly peptide reactive B cells and background noise, sequence clustering at different degrees of CDR3 amino acid sequence similarity was performed on either the full V-J gene sequence (Figure 23) or the CDR3 only (Figure 28). Four different types of clusters were defined. By "Peptide" clusters, we refer to clusters containing sequences originating from peptide positive cells only. "Control" clusters are composed of sequences from control cells only. "Mixed" clusters on the other hand contain a mixture of sequences derived from peptide positive and control cells. "CLL" clusters contain at least one of the sequences from the sorted $CD5^+$ CLL cells. Considering also V and J genes (Figure 23), the "Peptide" clusters did not substantially intermix with "Control" sequences. Even with more strict CDR3 identity thresholds, only a few peptide positive sequences clustered together, indicating that they originate from B cells with unique VJ pairings and CDR3 lengths. "Control" clusters were generally rare but became more frequent at lower CDR3 identity thresholds. The number of sequences belonging to mixed clusters increased most drastically under less stringent clustering conditions. The CLL sequences clustered stably together with no significant increase in members at lower CDR3 identity thresholds, indicating that they are unique and do not share similar sequences with other sorted cells. They clustered however with a subset of peptide sorted single B cell sequences from the same donor, confirming our previous suspicion of non-specific binding of peptide tetramers to the CLL cells. We found, that lowering the CDR3 identity threshold to below 85% was necessary to cluster single cell sequences into meaningful clones. Only when allowing this degree of mismatch were all the sequences belonging to CLL clones included into these clusters (explains transient mixed clusters at 100%-95% thresholds). Exclusion of some sequences at stricter thresholds is probably due to PCR and sequencing errors or alternatively could originate from biological variability.

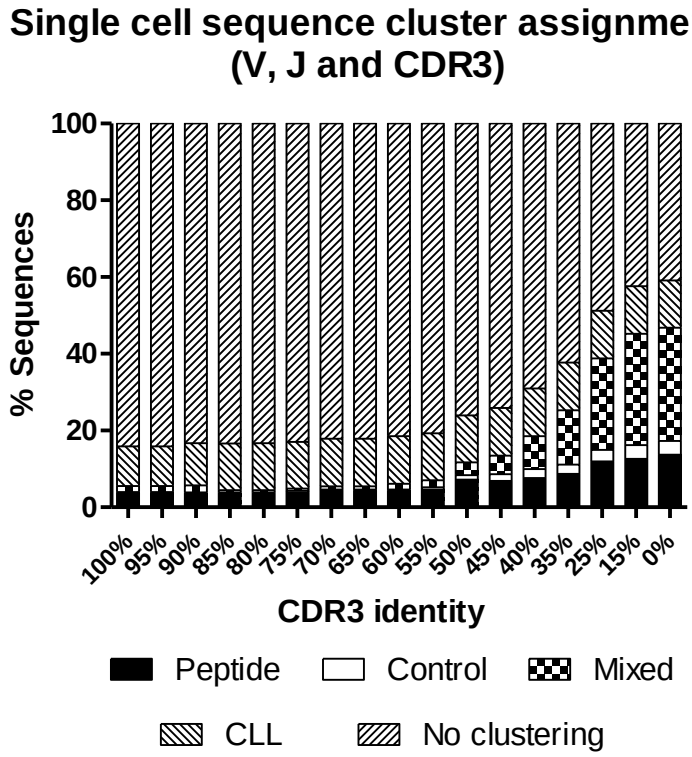


Figure 23: B cell receptor heavy chain sequences from peptide reactive and control cells were clustered according to same VJ gene usages and different CDR3 amino acid identities using the BcRep R package developed by Bischof et al. [2].

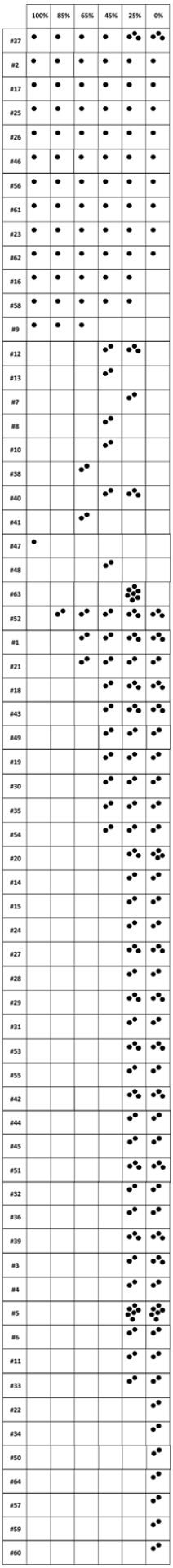


Figure 24: Representation of the evolution of individual "Peptide" clusters obtained with the BcRep R package [2] over different CDR3 identity thresholds. *Each row represents one "Peptide" cluster. Each box represents the composition of a cluster at the corresponding CDR3 identity threshold indicated at the top of each column. For this analysis, sequences sharing the same VJ gene assignments and the same CDR3 amino acid sequences were collapsed. Each dot represents a unique clustered sequence. This means, that one dot can be composed of several individual sequences. Several dots in one box means that sequences with different CDR3 amino acid sequences were assigned to belong to the same cluster. Clusters are only highlighted when they are solely composed of sequences from peptide reactive sorted B cells. An empty box means, that either sequences did not cluster together yet because of a too strict CDR3 identity threshold or that the cluster started to intermix with control sequences and therefore "disappeared" from the "Peptide" cluster pool. In the column on the left, clone identity numbers are represented. In the top row, the used CDR3 identity thresholds are indicated (100%, 85%, 65%, 45%, 25%, 0%).*

Next, we followed the evolution of individual “Peptide” clusters when reducing CDR3 identity thresholds (Figure 24). In this case, a “Peptide” cluster (n=64) is a cluster, that at least at one of the tested identity thresholds is composed of only peptide positive sequences. 16% (10/64) of the “Peptide” clusters were stable with no detectable intermixing with control sequences even at a CDR3 identity threshold of 0%. These clusters contained mostly two sequences with identical CDR3s. 5% (3/64) already clustered at 100% CDR3 identity but disappeared from the “Peptide” cluster pool, as they started to intermix with control sequences at CDR3 identity thresholds below 65% (1/64) or 25% (2/64). The vast majority of sequences (63%, 40/64) from “Peptide” clusters started clustering together at lower CDR3 identity (1/64 <100%, 2/64 <85%, 7/64 <65%, 23/64 <65%, 7/64 < 25%). Only 17% (11/64) of the “Peptide” clusters mixed with control sequences. This all confirms that peptide reactive B cells have heterogeneous and thus highly unique sequences.

At 65% CDR3 identity, 7% (44/626) of sequences from tetramer positive cells clustered together, but only 1% (4/353) of the randomly sorted cells (Figure 25). Sequences from cells of both groups intermixed to a similar extent, which occurred at 0.8% (Figure 25). These results indicate, that with our tetramer approach we are enriching for cells of a certain specificity. Sequences from clusters (65% CDR3 amino acid identity threshold) of seropositive donors were generally more distant from the germline (Figure 26) and from each other (Figure 27) as compared to the sequences – predominated by IgM isotype - isolated from the acute patients.

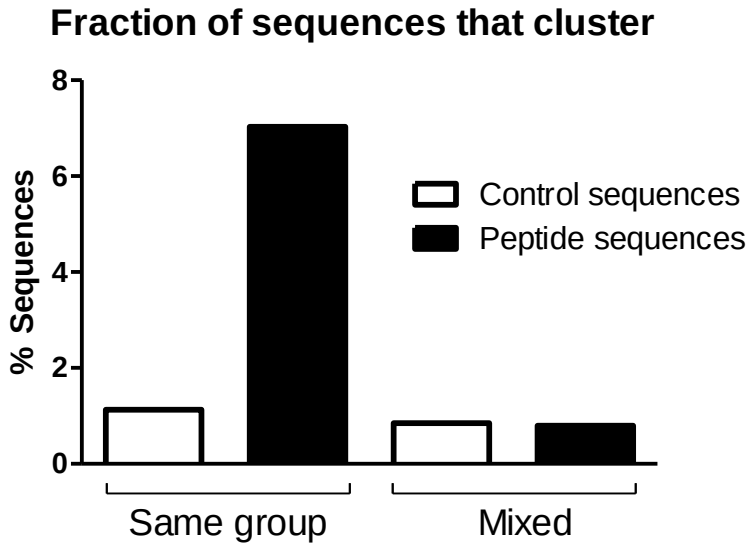


Figure 25: Percentage of single cell immunoglobulin heavy chain sequences that clustered with sequences from the same or the different group (control or peptide positive).

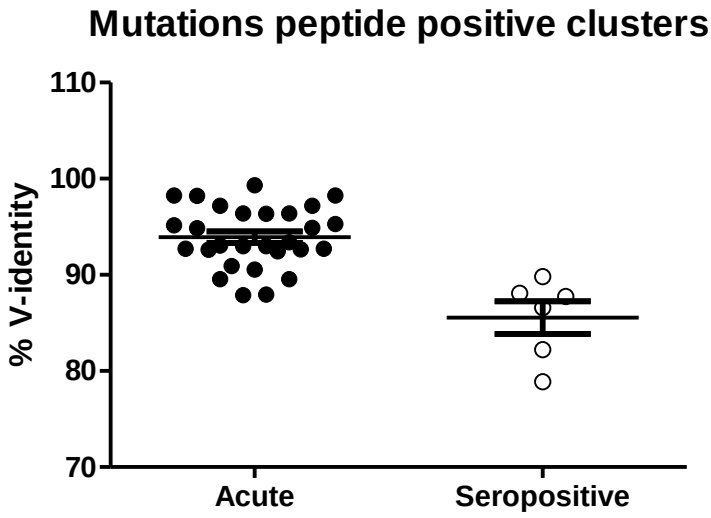


Figure 26: Mutation level in VlsE-C6 reactive sorted single cell immunoglobulin heavy chain sequences.

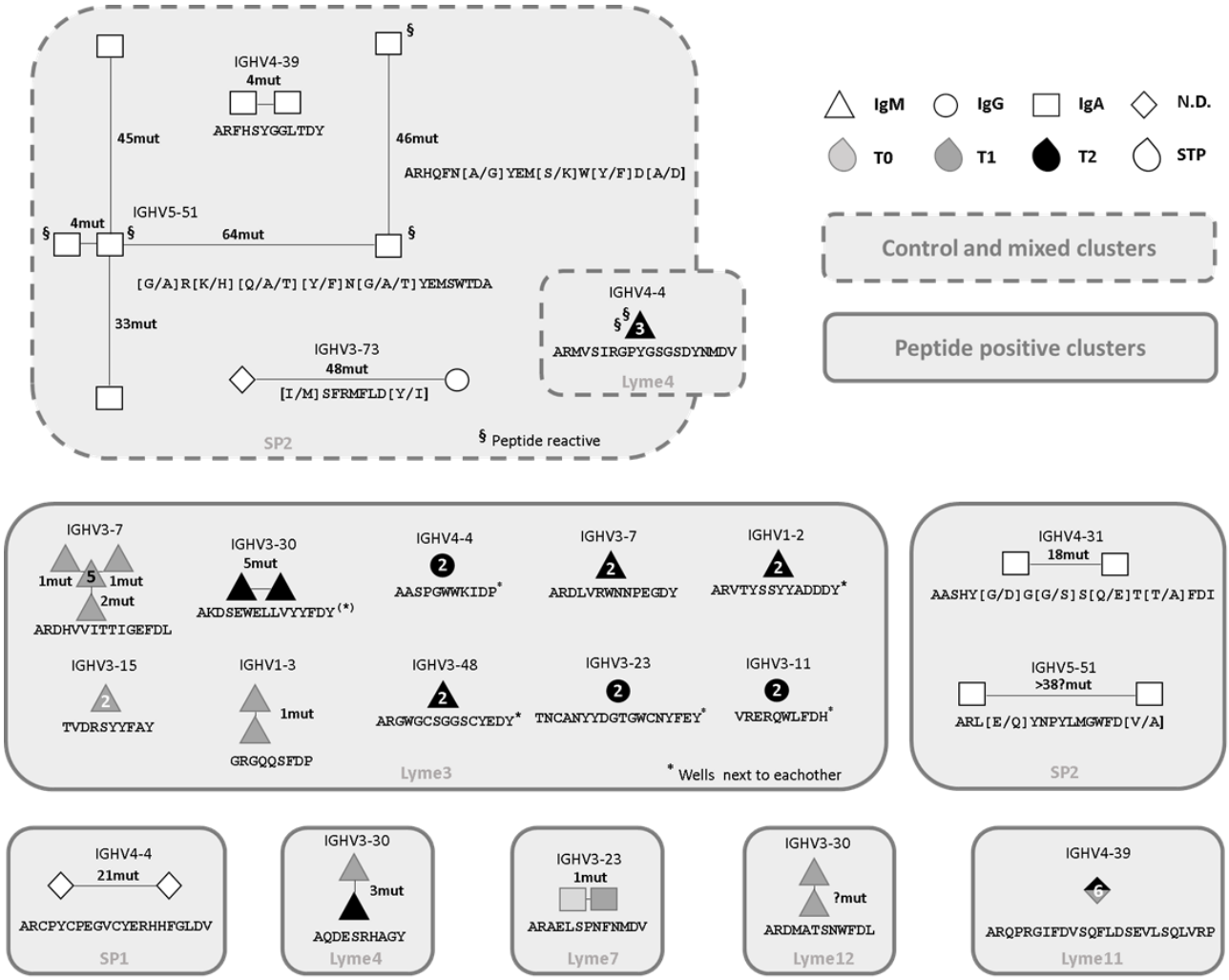


Figure 27: Representation of single cell sequence clusters. *Clusters obtained with the BcRep R package [2] at the 65% CDR3 identity threshold are depicted. CDR3 amino acid sequences including mismatched positions are indicated next to the corresponding cluster. Different timepoints are highlighted in different colors. Isotypes are depicted in different shapes. N.D.: isotype not determined because of incomplete sequence, STP: single timepoint measured only, mut: number of mutations separating the linked sequences. The length of each bar is proportional to the number of mutations between the sequences. In mixed clusters, sequences from peptide positive cells are marked with §. Some sequences were isolated from cells that were present in the wells next to each other (highlighted with *), which should be interpreted with caution.*

We also performed clustering based on CDR3 similarity alone, allowing a mismatch in every second (50% identity), third (67% identity), fourth (75% identity), fifth (80% identity), sixth (83% identity) or twelfth (92% identity) CDR3 amino acid (Figure 28). With this analysis, an overall higher clustering of sequences was observed, especially when allowing an increased number of CDR3 mismatches. Also in this case, sequences from peptide positive clusters stayed separated from control or CLL sequences and the majority of clusters responsible for an increased clustering at lower CDR3 identity contained a mix of peptide positive and control sequences.

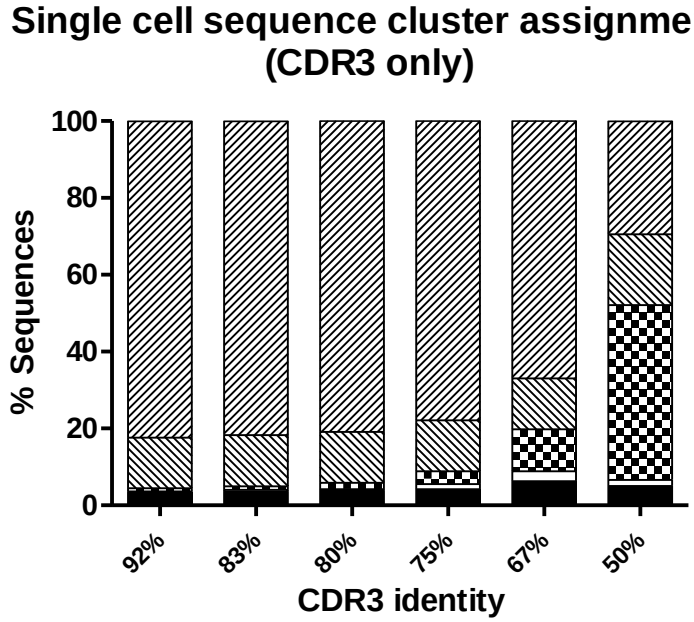


Figure 28: B cell receptor heavy chain sequences from peptide reactive and control cells were clustered considering only the CDR3 amino acid sequence using the python script from Galson et al. [12,127,171]. *The same color code as for Figure 23 was used.*

4.5 Closer analysis of mutations in immunoglobulin heavy chain sequences from CD5$^+$ B cell subpopulation of Lyme11

42 of the sorted single CD5$^+$ B cells from Lyme11 could be assigned to three different clones. Two of those were quite distant from the germline (clone1: 91% and clone2: 90% V-region identity) indicating that they might have participated in germinal center responses. The third and less prominent clone was less mutated (clone3: 96% V-region identity). The sequences of these three clones were not related to any one of the previously described major stereotypic CLL B cell receptor subsets [361,362].

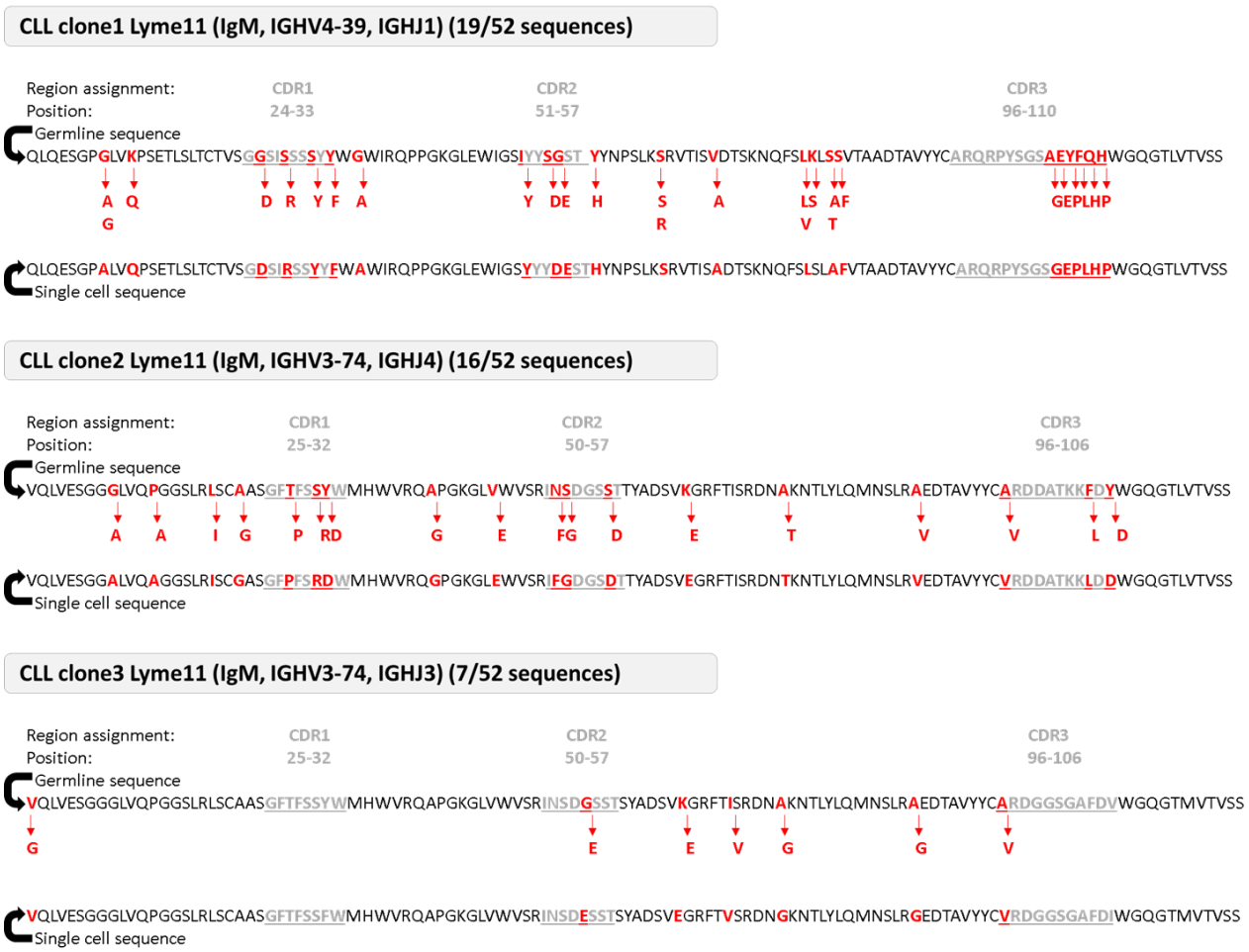


Figure 29: Closer analysis of mutations occuring in dominant clones from single sorted CD5$^+$ B cells from Lyme11.

4.6 B cell repertoire analysis with next generation sequencing

Next we investigated changes in the bulk B cell repertoire induced by *Borrelia* and the tick bite. We applied a barcoding technique similar to the one from Vollmers et al. [1] which we adapted to the Ion Torrent next generation sequencing platform (Figure 30, Figure 31A). First, we compared sequencing replicates, library replicates and biological replicates (Figure 31B). 97±1% of CDR3 sequences overlapped between IgG sequencing replicates. IgG library replicates still overlapped by 74±24%. Biological replicates overlapped by 32±11% for IgG and 17±10% for IgM. After data cleaning and consensus sequence building with an in-house pipeline developed by colleagues, an average number of 24012±39051 (median: 9470; min: 594, max: 218302) productive sequences were obtained for IgG samples and 83914±65861 (median: 66878; min: 1727; max: 231257) for IgM. Similar to the

randomly sorted single cells described above, the most abundant V and J genes were IGHV3-23, IGHV3-30, IGHV3-7, IGHJ4 and IGHJ6 (Figure 32, Figure 33). A considerable (on average 8-30% ranging from 0.31-59% for IgM and 6-12% ranging from 0-35% for IgG) CDR3 amino acid sequence overlap was observed between the different timepoints of the same individuals (Figure 31C,D). Interestingly the IgM overlap was on average two to three times higher in acute patients and tick bite donors compared to healthy individuals (Figure 31C). The high overlap in the tick bite group was only present in two of the three donors analyzed. In Tick1 we observed one largely expanded clone that predominated all three samples (14%(T0), 10%(T1), 13%(T2) of the IgM repertoire) while in Tick2 three such largely expanded clones (17%, 6%, 5% (T0), 13%, 9%, 4% (T1), 11%, 11%, 3% (T2) IgM repertoire) have been observed. For IgG, the average sequence overlap between different timepoints was also higher in acute patients than in the case of tick bite and healthy controls (Figure 31D). As expected [12,14], interdonor overlap was very small (Figure 31E,F), indicating that each donor has a fairly unique B cell receptor repertoire. IgM CDR3 amino acid sequence overlap occurred ~1.5 times more often in healthy individuals than in the other groups (Figure 31E). There was eventually no overlap between IgG samples from different donors (Figure 31F). We observed a tendency towards a higher concentration of final IgG libraries and number of IgG reads for acute Lyme disease samples (Figure 34F,H).

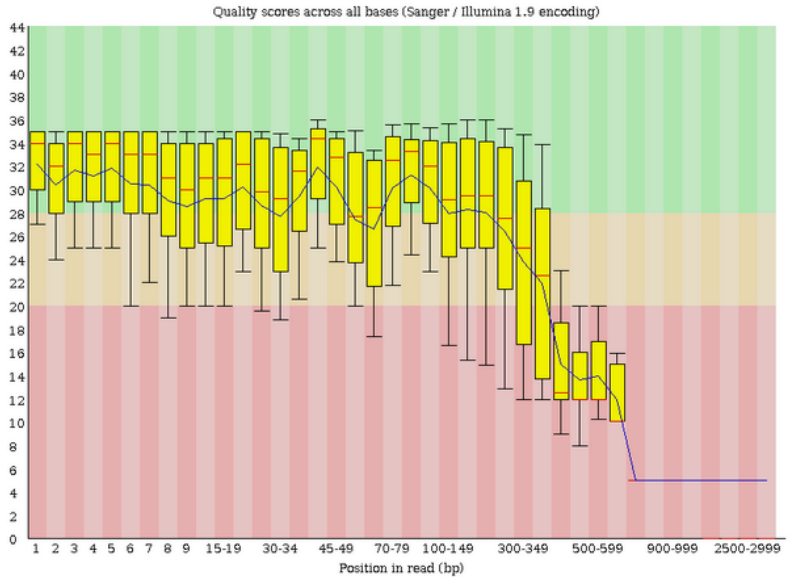


Figure 30: Representative quality scores (y axis) at the different positions within the sequences (x axis). *Due to sequence quality drop off at the end of Ion Torrent reads, we decided to change from the double sided UID method described by Vollmers et al [1] for the Illumina platform to a single sided UID method. In this way a high quality of barcodes could be guaranteed by placing them at the beginning of the sequences. After correct assignment of UIDs into groups, the remaining bases could then be corrected by sequence alignment.*

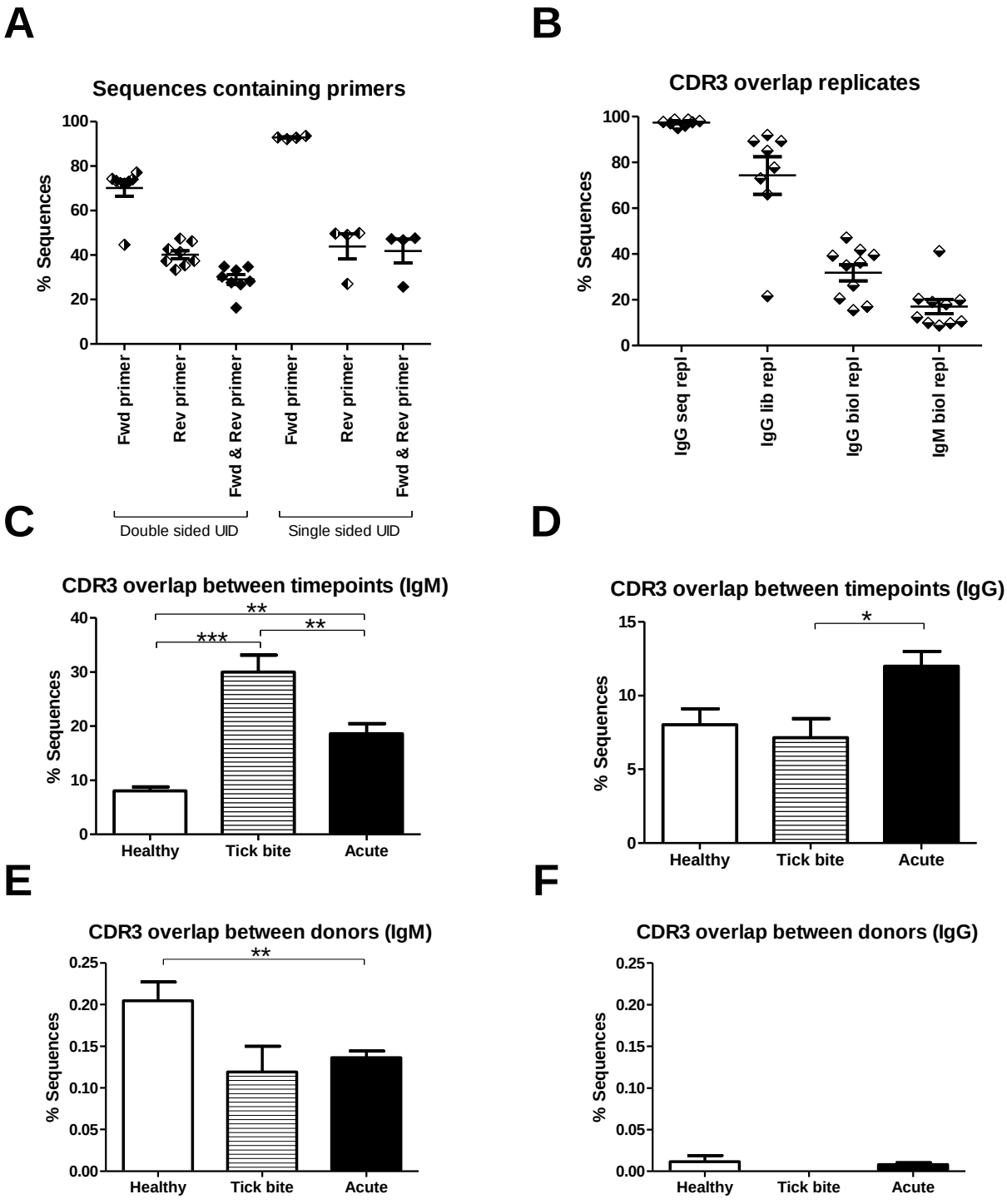


Figure 31: Setup of the deep sequencing method to analyze patients' B cell repertoires. *(A) The percentage of reads containing forward and reverse primer sequences demonstrates the extensive loss of reads due to quality drop off at the end of sequences described in Figure 30. (B) The extent of CDR3 amino acid sequence overlap between sequencing replicates (seq repl), library replicates (lib*

repl) and biological replicates (biol repl) are represented as percentage of final consensus sequences. CDR3 amino acid sequence overlap was also determined between samples from different timepoints (C,D) and donors (E,F). Groups were compared with One-way Analysis of Variance test followed by Tukey's Multiple Comparison test.

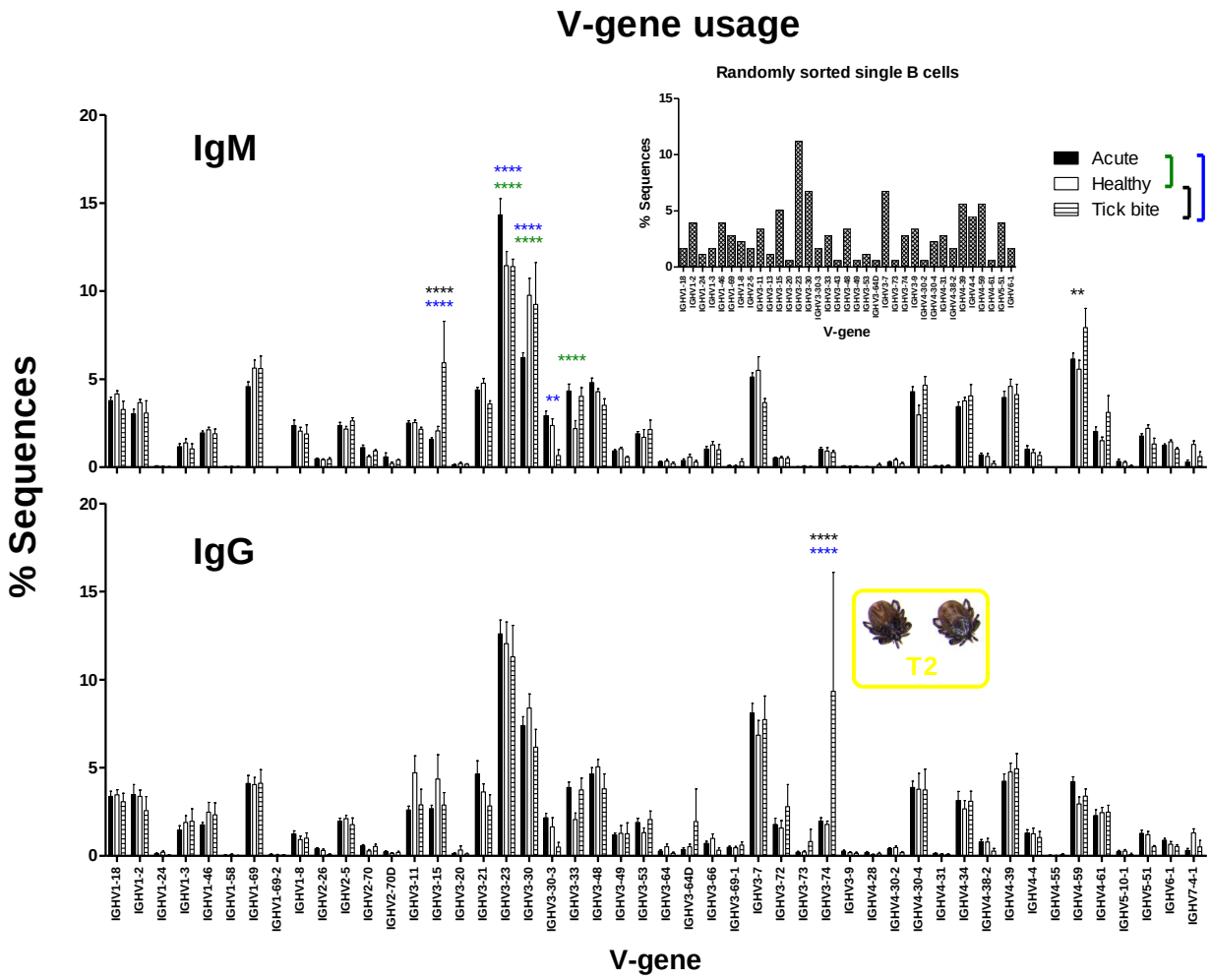


Figure 32: V-gene distribution of next generation sequencing samples compared to that one of sequences from randomly sorted single cells. *Two Way ANOVA followed by Bonferroni Multiple Comparisons tests were used to compare the groups.*

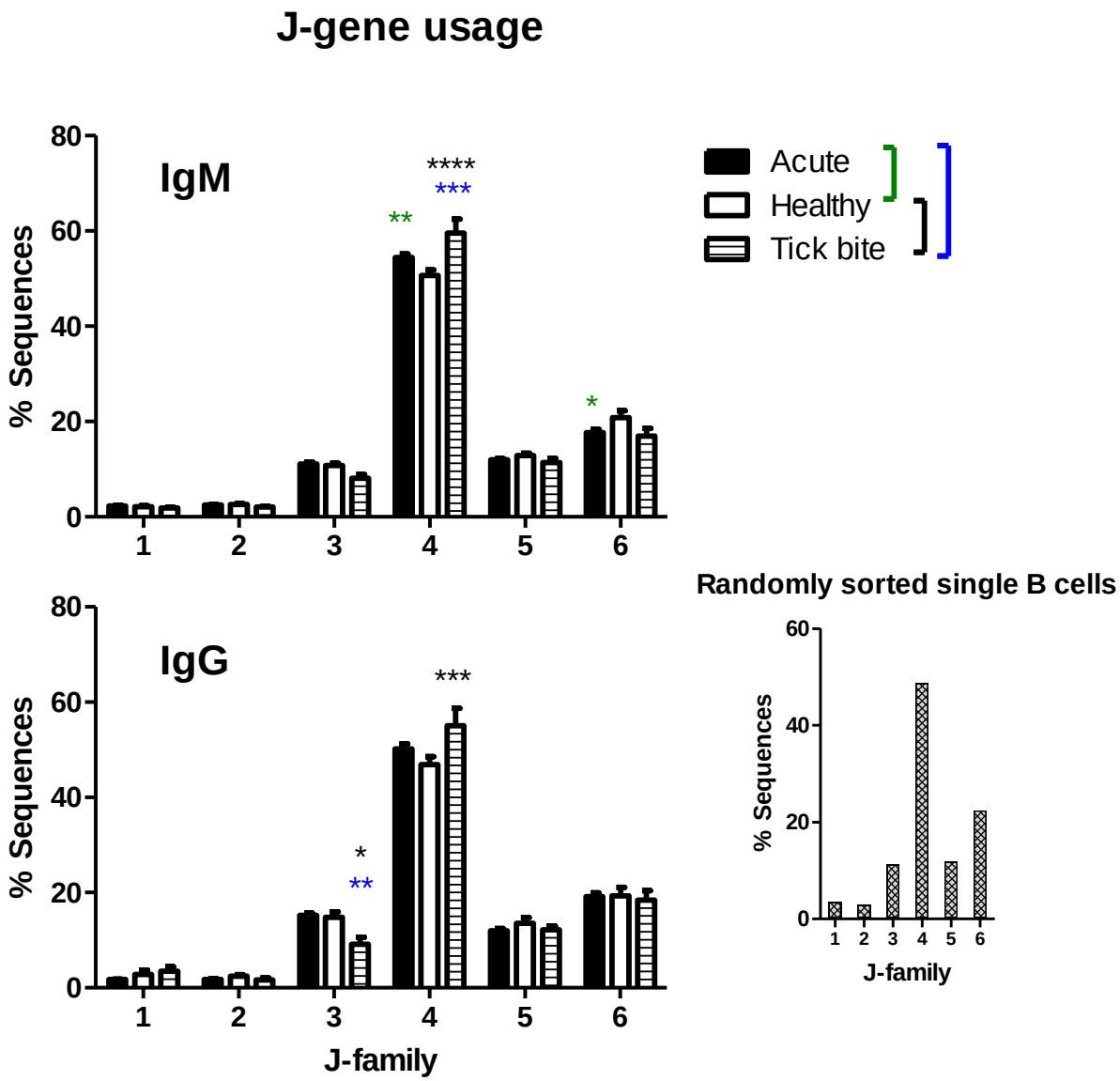


Figure 33: J-gene distribution of next generation sequencing samples compared to that one of sequences from randomly sorted single cells. *Two Way ANOVA followed by Bonferroni Multiple Comparisons tests were used to compare the groups.*

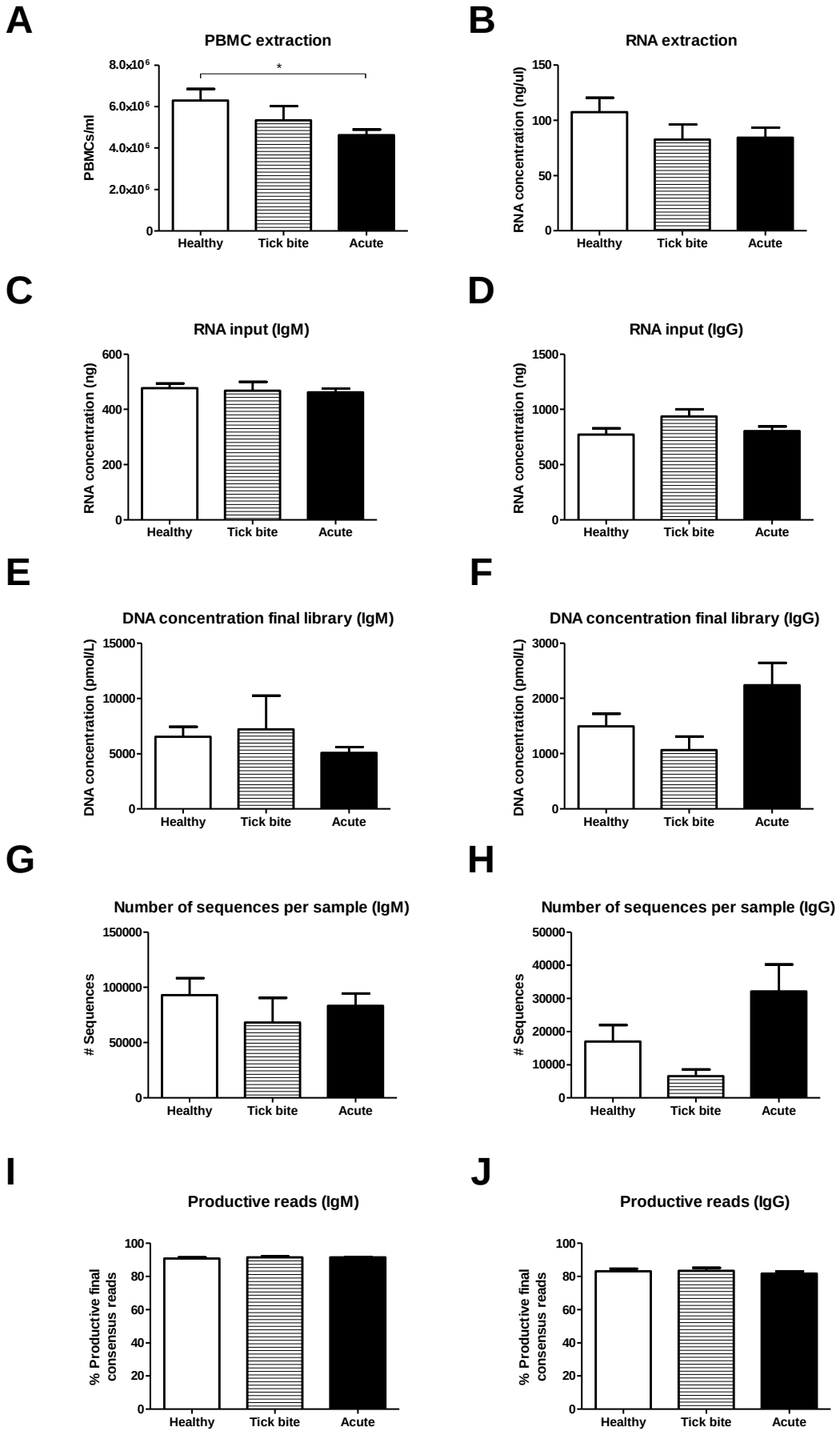

A
PBMC extraction
PBMCs/ml
8.0x10⁶
6.0x10⁶
4.0x10⁶
2.0x10⁶
0
*
Healthy
Tick bite
Acute
B
RNA extraction
RNA concentration (ng/ul)
150
100
50
0
Healthy
Tick bite
Acute
C
RNA input (IgM)
RNA concentration (ng)
600
400
200
0
Healthy
Tick bite
Acute
D
RNA input (IgG)
RNA concentration (ng)
1500
1000
500
0
Healthy
Tick bite
Acute
E
DNA concentration final library (IgM)
DNA concentration (pmol/L)
15000
10000
5000
0
Healthy
Tick bite
Acute
F
DNA concentration final library (IgG)
DNA concentration (pmol/L)
3000
2000
1000
0
Healthy
Tick bite
Acute
G
Number of sequences per sample (IgM)
Sequences
150000
100000
50000
0
Healthy
Tick bite
Acute
H
Number of sequences per sample (IgG)
Sequences
50000
40000
30000
20000
10000
0
Healthy
Tick bite
Acute
I
Productive reads (IgM)
% Productive final consensus reads
100
80
60
40
20
0
Healthy
Tick bite
Acute
J
Productive reads (IgG)
% Productive final consensus reads
100
80
60
40
20
0
Healthy
Tick bite
Acute

Figure 34: Control for equal library preparation between samples from different groups. *Comparison of PBMC (A) and RNA (B) extractions, RNA input (C,D), DNA concentrations of final libraries (E,F), final number of consensus sequences (G,H) and the percentage thereof that was productive (I,J) between groups using a One-way Analysis of Variance test followed by Tukey's Multiple Comparison test.*

Because of the uniqueness of the repertoire sequences of each donor, inter-individual sample overlap was also assessed using a less stringent approach. Sequences from the two groups (controls and acute patients) were clustered separately using the python script from Galson et al. [12,127,171]. Different CDR3 identity thresholds once taking V and J gene information into account (Figure 35A) and once only considering CDR3 sequences (Figure 35B) were tested. Similar trends were observed at all the tested conditions. Also with this approach, a slightly higher IgM CDR3 sequence overlap could be observed among healthy individuals as compared to the two other groups (Figure 31E,Figure 35B). Furthermore an increased overlap between acute Lyme disease patients' IgG repertoires was apparent. At 83% CDR3 identity including V and J gene information (Figure 35A) we observed that one largely expanded clone detected in all samples from Tick1 overlapped with all the other samples from another tick bite donor (Tick3). This high overlap was most striking under these clustering conditions. Under the same conditions we also observed a larger number of acute patient samples that shared an unusual high frequency of reads with other samples from that same group. In that case it was samples from different donors that shared such a high fraction of sequences with one of the other samples from the group. When we calculated the median from these individual values for each donor, a significant difference between acute Lyme disease patients and controls was observed only for IgG (Figure 36).

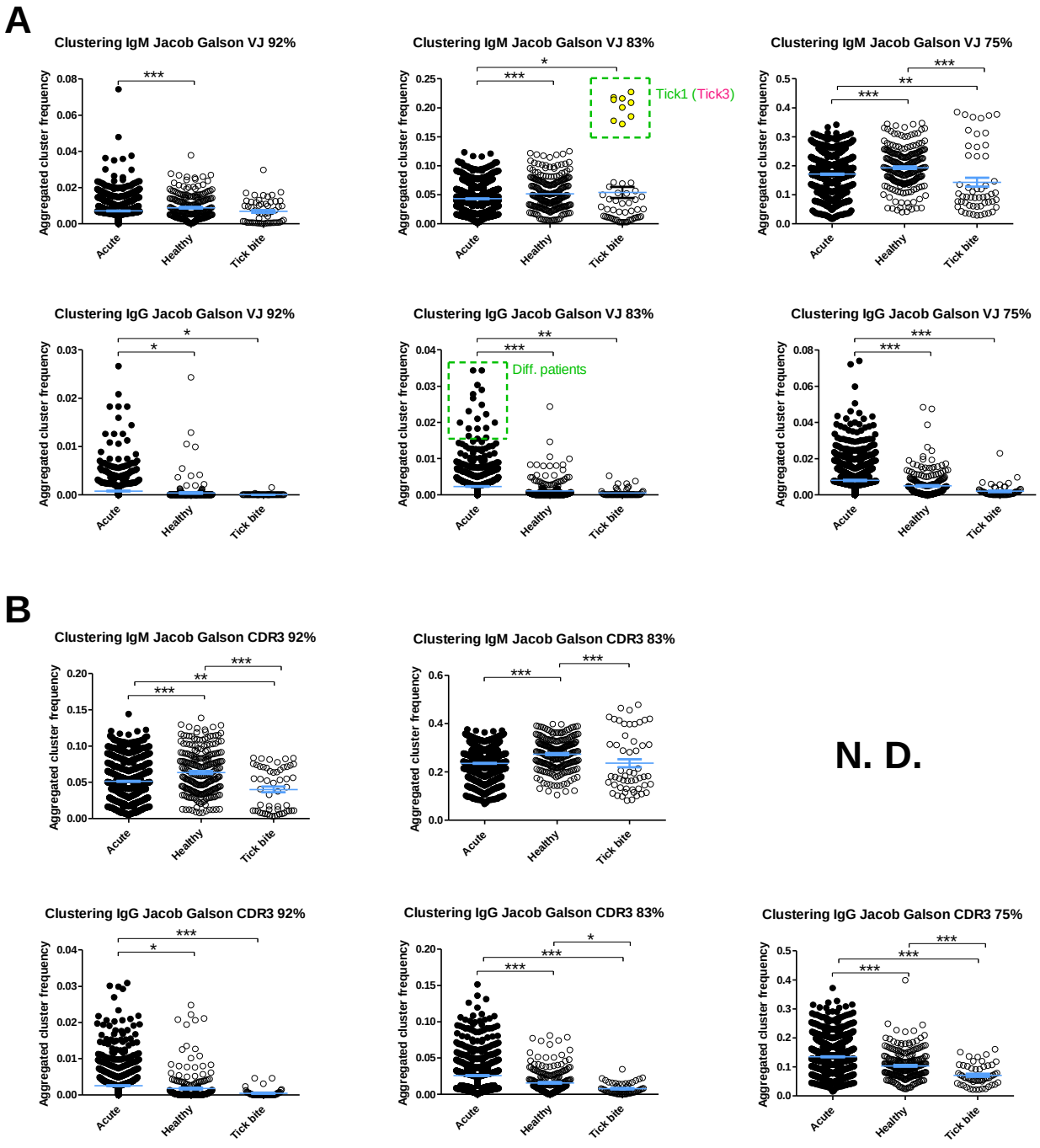


Figure 35: Cluster overlap between samples at different CDR3 identity thresholds. *Sequences from acute donors' and control samples were clustered separately with the python script from Galson et al [12,127,171] once taking V and J gene information into account (A) and once only considering CDR3 sequences (B). Groups were compared using a One-Way Analysis of Variance followed by Tukey's Multiple Comparison Test. Tick1(Tick3): One large cluster from Tick1 that was present in all three timepoints was found to cluster together with a few sequences from Tick3. Diff. patients: values are from different patients. N.D.: Not determined..*

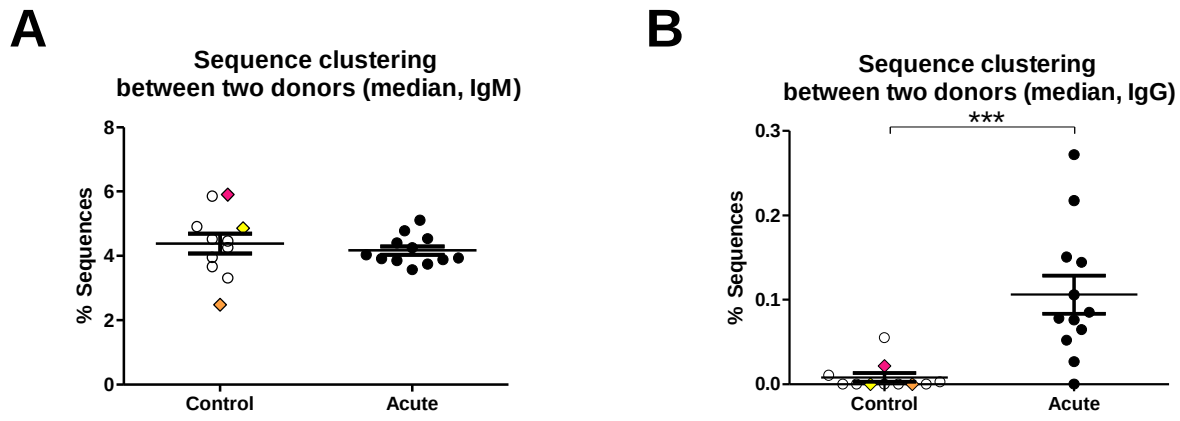


Figure 36: Assessment of sequence cluster overlap between different donors. *Sequences were clustered based on whether they contain the same V and J genes as well as similar CDR3 amino acid sequences (one in six mismatches) using the python script from Galson et al. [12,127,171]. For each donor, the percentage sequence overlap between each sample from that donor with individual samples from all the other donors of that group was determined (Figure 35) and the median from all these values calculated. Data points corresponding to the three controls with a recent tick bite are highlighted with rhombi colored as in the other graphs. A Two-tailed Unpaired t-test was used to compare values from the two groups.*

Since a recent immune response should affect both, the IgM and IgG levels, we tested whether these parameters for IgM and IgG correlate with each other. When assessing the percentage of the repertoire (with all samples from the same participant pooled) that clusters containing a mix of IgM and IgG isotypes are making up, only 20% (1 of 5) of healthy donors but 67% (8 of 12) of acute Lyme disease patients showed elevated levels (Figure 37A). Interestingly two of the three tick bite donors that also showed largely skewed IgM repertoires (Figure 31C) also contained very high levels of sequences belonging to mixed clusters indicating that there is something going on at both IgG and IgM levels. When plotting IgM versus IgG sequence overlap between samples from different timepoints, a weak but positive correlation was seen in the patient group while a negative correlation was observed for controls (Figure 37B). The latter was mainly due to a strong perturbation at the IgM level in the absence of a strong effect on the IgG repertoire in the two tick bite donors. The four donors for which we observed a change in VlsE-C6 IgG titers over the sampling period with our in-house peptide ELISA (Lyme8 & Lyme2↑, Lyme 4 and Lyme13↓, Lyme1↑↓, Table 2) also showed the highest IgG CDR3 amino acid sequence overlap between the different analyzed timepoints. When assessing the degree of correlation between sequence overlap among timepoints and interdonor sequence clustering at the IgG level, we observed an anticorrelation for acute patient samples (Figure 37C), indicating

that expanded clones probably have a lower chance to overlap between donors as compared to smaller ones.

In order to exclude, that the differences in clustering between groups (Figure 36B) were due to different numbers of samples (the higher the number of patients compared, the higher the probability that sequences overlap), we clustered IgG samples from all donors and as in Figure 36 assessed for each donor the degree of overlap with donors from all the other groups. This showed, that acute patients have higher overlaps only with members of their own group (Figure 37D). However, both in the patient and control group, the majority of clusters overlapped only between two donors (Figure 37E) and the majority of clusters overlapping between a higher number of acute patients was also present in healthy donors (Figure 37F).

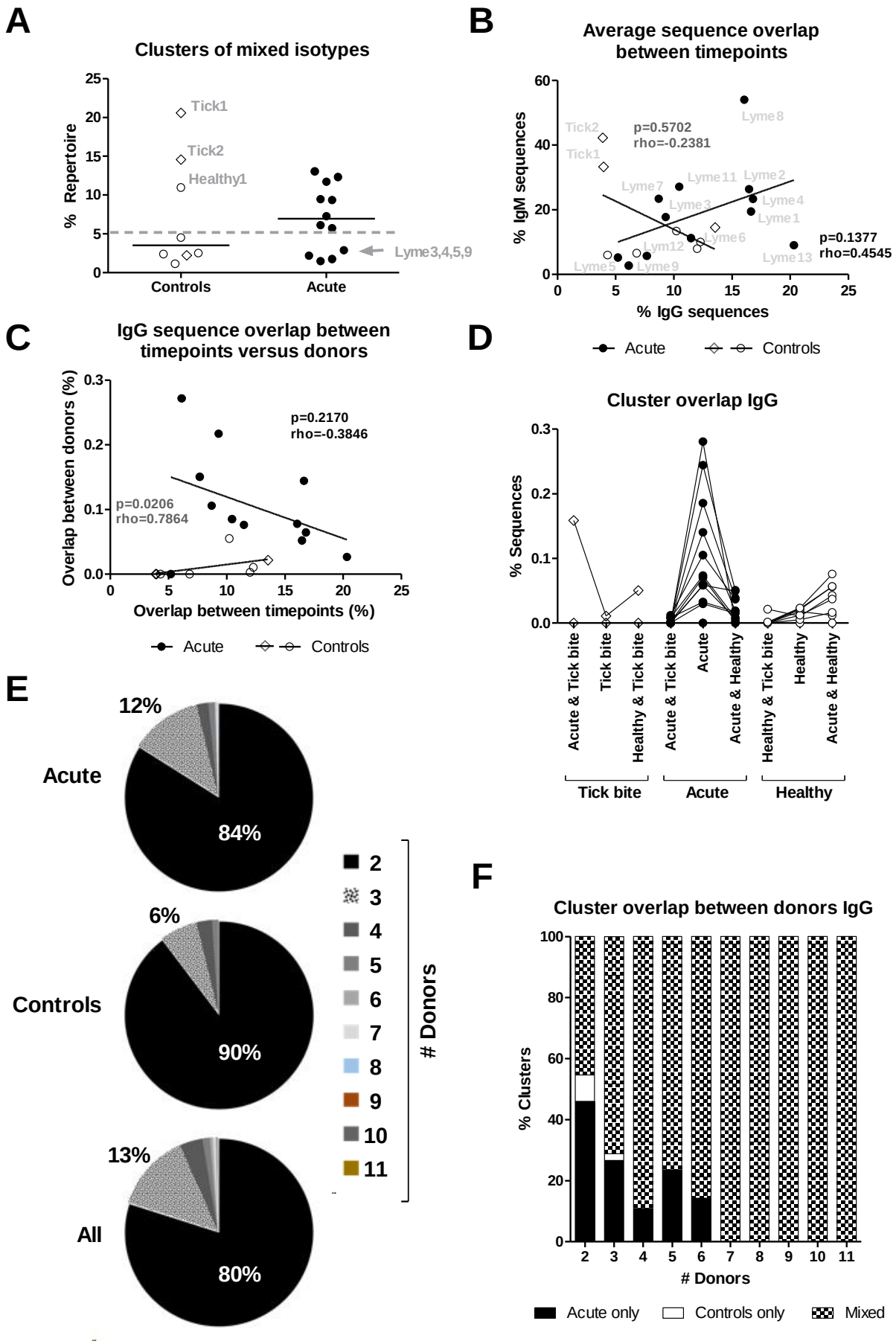
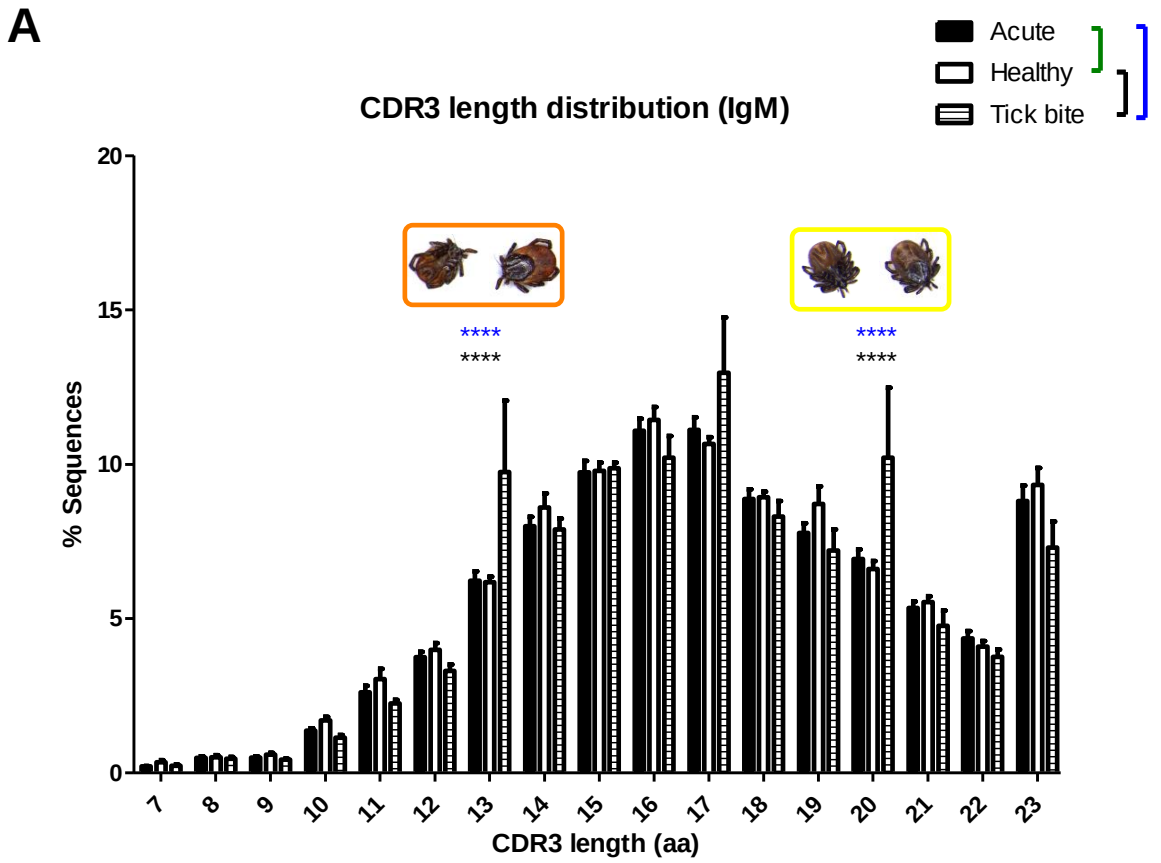

A
Clusters of mixed isotypes
% Repertoire
Tick1
Tick2
Healthy1
Lyme3,4,5,9
Controls
Acute
B
Average sequence overlap between timepoints
% IgM sequences
% IgG sequences
p=0.5702
rho=-0.2381
p=0.1377
rho=0.4545
Acute
Controls
C
IgG sequence overlap between timepoints versus donors
Overlap between donors (%)
Overlap between timepoints (%)
p=0.2170
rho=-0.3846
p=0.0206
rho=0.7864
Acute
Controls
D
Cluster overlap IgG
% Sequences
Acute & Tick bite
Tick bite
Healthy & Tick bite
Acute & Tick bite
Acute
Acute & Healthy
Healthy & Tick bite
Healthy
Acute & Healthy
Tick bite
Acute
Healthy
E
Acute
12%
84%
Controls
6%
90%
All
13%
80%
Donors
F
Cluster overlap between donors IgG
% Clusters
# Donors
Acute only
Controls only
Mixed

Figure 37: Closer analysis of sequence clustering. *(A) Clusters obtained using Change-o [338] that contained sequences of both IgM and IgG isotypes were selected and the percentage total repertoire (pooled samples as before) they make up is represented. (B) Correlation between average IgM and IgG CDR3 amino acid sequence overlap between timepoints. (C) Correlation between IgG CDR3 amino acid sequence overlap between timepoints (average) and clustering between donors (median) [12,127,171] is represented. Spearman's Rank Correlation Hypothesis Testing was done manually in Excel. Sequence clustering between all donors was determined using the python script from Galson et al. [12,127,171]. Median values were calculated as before (Figure 36) but this time also assessing sequence overlap between samples belonging to different groups. Lines connect values from the same donor. (E,F) Distribution of overlapping clusters among donors.*

4.7 Comparison of different B cell repertoire parameters among groups

In order to determine the degree of repertoire distortion induced by *Borrelia* infection, we first compared CDR3 amino acid length distribution (spectratypes) between samples from the different groups (Figure 38). The strong perturbations observed in the tick bite group could be explained by expansions in repertoires from individual donors rather than common changes observed in all the samples from that group. Since at this level of analysis, expansions of different clones (having different CDR3 lengths) in different patients could mask a potential distortion of repertoires, we also determined diversity based on nucleotide or amino acid sequences (Figure 39). Surprisingly, with different indices we repeatedly observed a slight increase in repertoire diversity in acute Lyme disease patient IgG samples as compared to the controls (Figure 39B,F&H). Overall differences between biological replicates were however large. Even when performing hierarchical clustering using diversity profiles of individual samples [3] (which should be less dependent on sequencing depth) biological replicates did generally not cluster together (Figure 40).


A
Acute
Healthy
Tick bite
CDR3 length distribution (IgM)
****
****
****
****
% Sequences
0
5
10
15
20
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
CDR3 length (aa)

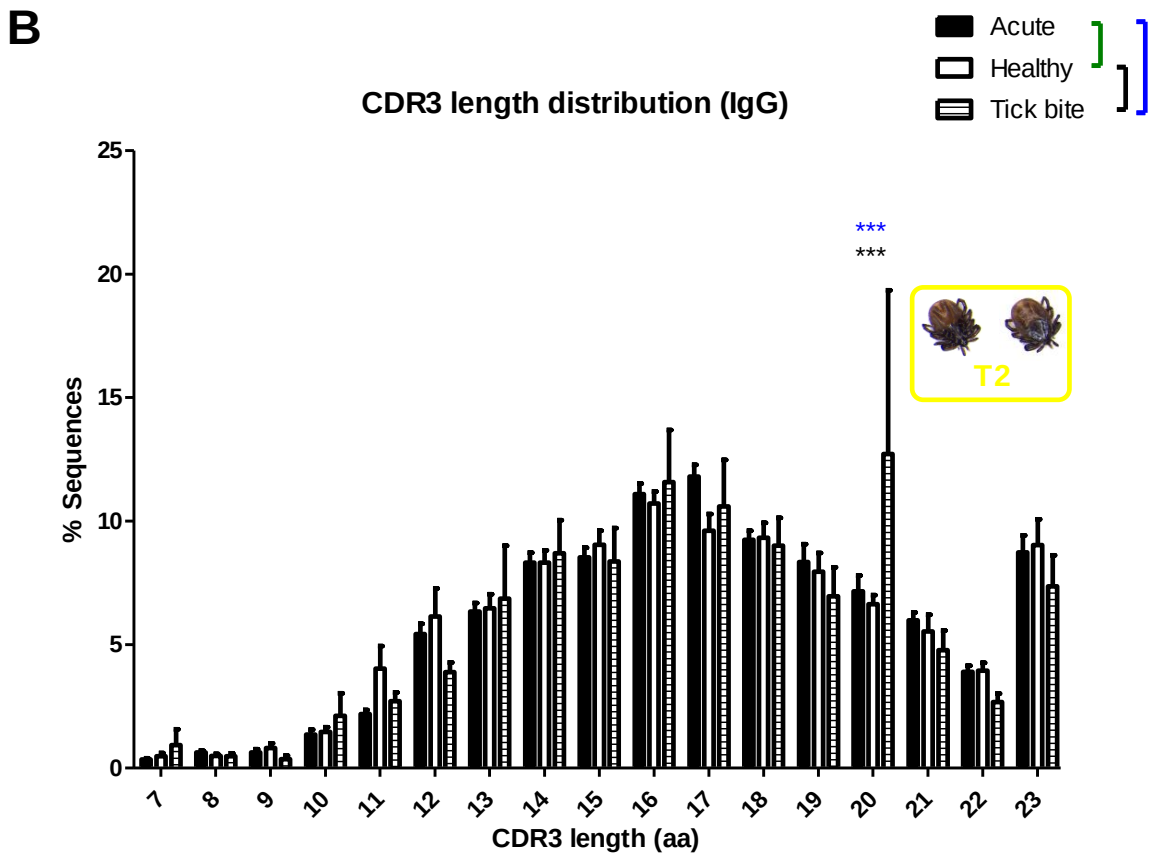

B
Acute
Healthy
Tick bite
CDR3 length distribution (IgG)
***
***
T2
% Sequences
0
5
10
15
20
25
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
CDR3 length (aa)

Figure 38: CDR3 amino acid spectratype analysis. *VDJtools [348] was used to extract CDR3 amino acid length distributions from deep sequencing data. Two-way RM ANOVA followed by Bonferroni multiple comparisons test was used to compare the groups. The tick bite donors (yellow: Tick1; orange: Tick2) responsible for the observed distortions are highlighted above the expanded peaks. T2: timepoint 2.*

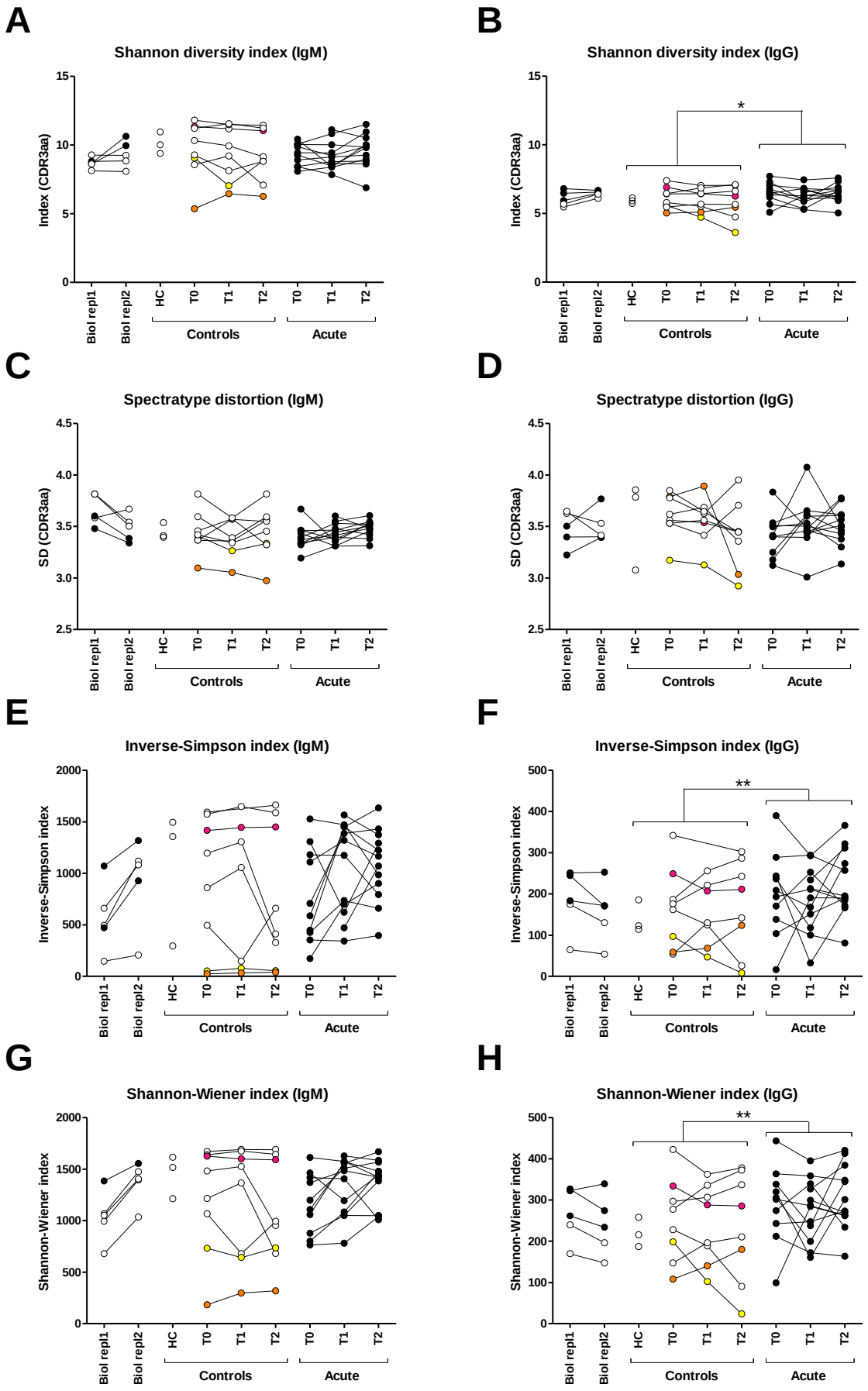

A
Shannon diversity index (IgM)
Index (CDR3aa)
B
Shannon diversity index (IgG)
Index (CDR3aa)
*
C
Spectratype distortion (IgM)
SD (CDR3aa)
D
Spectratype distortion (IgG)
SD (CDR3aa)
E
Inverse-Simpson index (IgM)
Inverse-Simpson index
F
Inverse-Simpson index (IgG)
Inverse-Simpson index
**
G
Shannon-Wiener index (IgM)
Shannon-Wiener index
H
Shannon-Wiener index (IgG)
Shannon-Wiener index
**
Biol repl1
Biol repl2
HC
T0
T1
T2
Controls
Acute

Figure 39: Changes in B cell repertoire diversity upon *Borrelia* infection. *Different diversity indices were compared. (A-B) Shannon-diversity indices were calculated manually using uncollapsed CDR3 amino acid sequences (C-D) Spectratype distortions were determined from deep sequencing data as in [20]. (E-G) Inverse-Simpson and Shannon-Wiener diversity indices were calculated with VDJtools [348]. Samples from controls and acute patients were pooled and differences between these two groups compared using an Unpaired t test.*

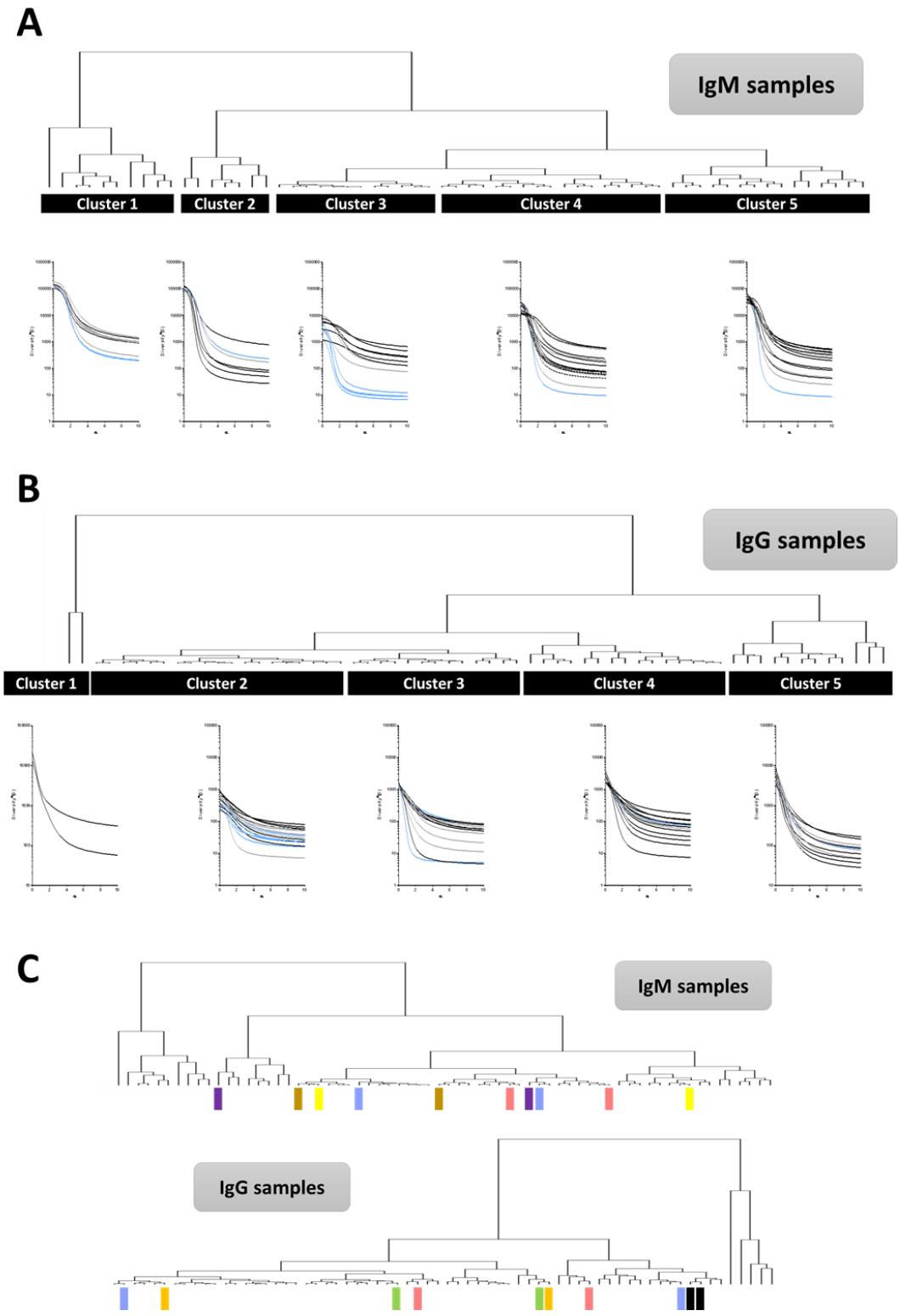


Figure 40: Hierarchical clustering of diversity profiles[3]. *(A-B) Diversity profiles of individual samples belonging to the highlighted clusters are represented in the graphs. Blue curves: tick bite donors. Grey curves: healthy donors. Black curves: acute donors. y axis: Diversity ($^{\alpha}D$). x axis: α. (C) Clustering was reperformed this time also including biological replicates. Biological replicates belonging to the same donor are highlighted in the same color at the bottom of each tree.*

Since recent mouse studies suggested that *Borrelia* actively influence the B cell immune response [289-293,303], we tested whether a similar effect also occurs in humans. To understand whether germinal center responses might be affected in acute Lyme disease patients, we assessed mutation frequencies of B cells. We were able to confirm, that sequences from IgM repertoires are less mutated as compared to those from IgG samples (Figure 41). Differences between the three groups (Acute, Healthy and Tick bite) could be observed at the 94% germline identity peak in IgM. When looking at the profile of the curves, healthy individuals' sequences were the least mutated, acute Lyme disease patients' repertoires showed an intermediate mutation frequency, while the tick bite individuals contained most mutated sequences within their "distorted" IgM repertoire (Figure 41, top). Interestingly, in acute Lyme disease patients, we found a higher number of IgG sequences with lower mutation frequencies as compared to control repertoires, while the opposite was true for individuals which were stung by *Borrelia* negative ticks (Figure 41, bottom).

It was a bit surprising to see that with our improved UID barcoding technique (that should remove the majority of sequencing errors) the major germline IgM peak lies at 94% instead of 100% V-region identity. Also when comparing the V-identity distribution with other samples from our laboratory for which no UID barcoding technique was used, we observed that the curves from these samples were less distant from the germline than those of the samples from the Lyme disease study. In order to understand what happened, we picked out a few sequences from the largely mutated peak from Tick2 for closer analysis and found out, that sequences still contained the P adapter and some nucleotides hat were probably added by the Ion Torrent machine at the end of the sequencing runs. Although the V-gene seemed to be correctly assigned, the V-identity values are not correct as IMGT/High V-QUEST [354] seemed to have included this non-V-gene sequence into the calculations. Since similar nucleotides seem to be added at the end of each sequence, we expect a simple shift of the whole curve without individual sequences being affected too much. In order to verify that we are indeed removing errors with our approach, we compared V-identity distribution and productive reads before and after consensus sequence building. For this, we randomly chose two Chips (one IgM and one IgG) and extracted the sequence with the best quality from each UID to compare this dataset to the final analyzed reads of the same samples. These data are indicative of a substantial reduction in mutations and increase in quality of the reads after sequence cleanup (V-identity before cleanup: 83.22±0.40 (IgM), 79.24±0.83 (IgG), after cleanup: 89.97±2.90 (IgM), 84.38±0.82 (IgG); Number of productive reads before cleanup: 85.95±0.95 (IgM), 76.29±5.87 (IgG), after cleanup: 93.13±0.74 (IgM), 87.79±1.28 (IgG).

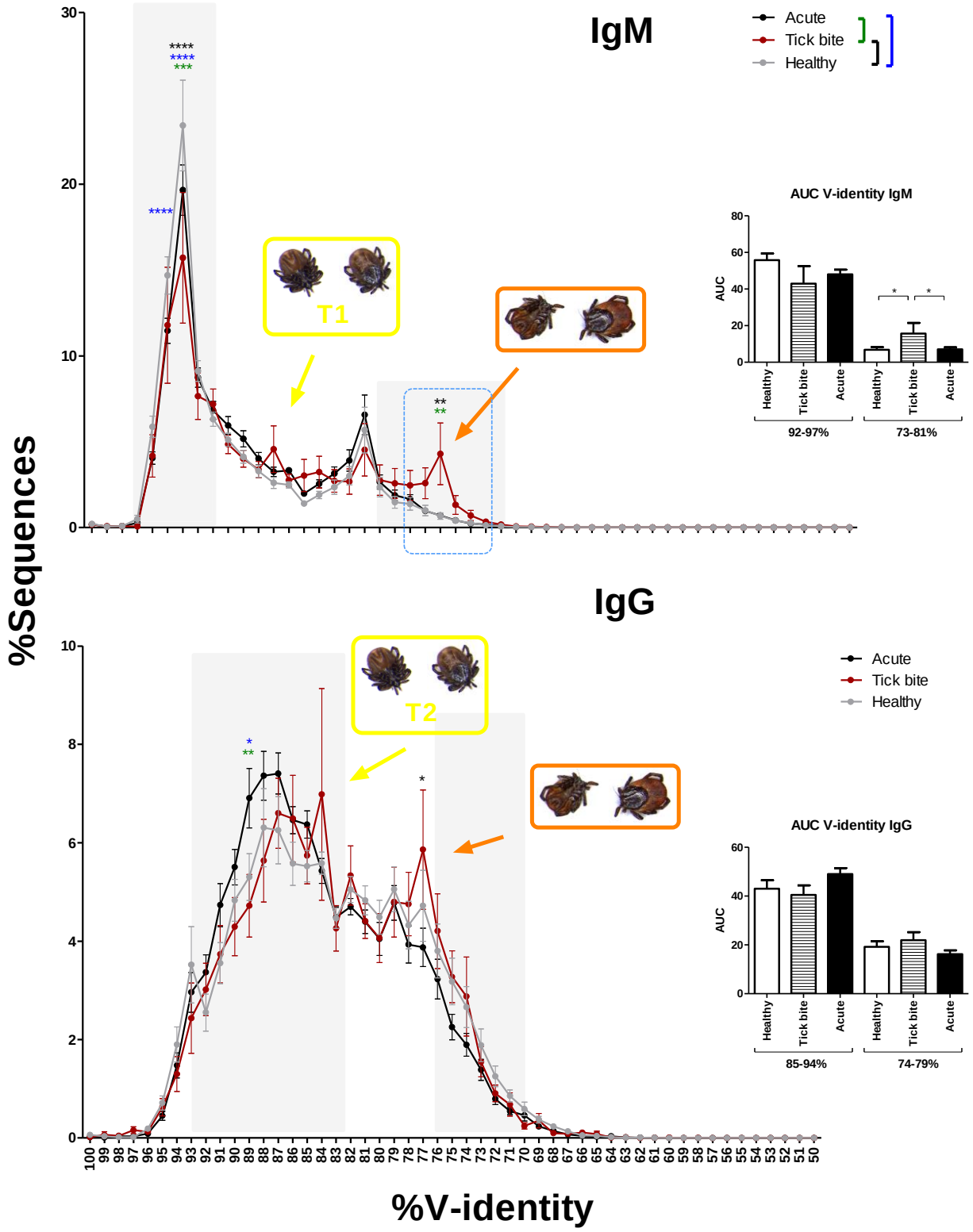


Figure 41: V-identity distribution. *IMGT/High V-QUEST [354] was used to determine distance of reads from the closest germline gene for IgM (top) and IgG (bottom) repertoire data. The highly mutated IgM and IgG peaks were mainly due to one of the three individuals with a recent tick bite*

(orange: Tick2) while the other less mutated peaks were mainly due to the sample from the second timepoint (T1 in IgM) or the third timepoint (T2 in IgG) from another individual with a recent tick bite (yellow: Tick1). Groups were compared with Two-way Analysis of Variance followed by Bonferroni Multiple Comparisons Tests. On the top right of each graph the calculations of areas under the curve of the highlighted (in grey) regions are depicted. Values of samples were compared using a One-way Analysis of Variance followed by Turkey's Multiple Comparisons test. Differences between groups are hardy detectable with this method. Individual sequences from the unusually high mutated peak from Tick2 (highlighted with a blue dashed line) were more closely analyzed (Figure 42).

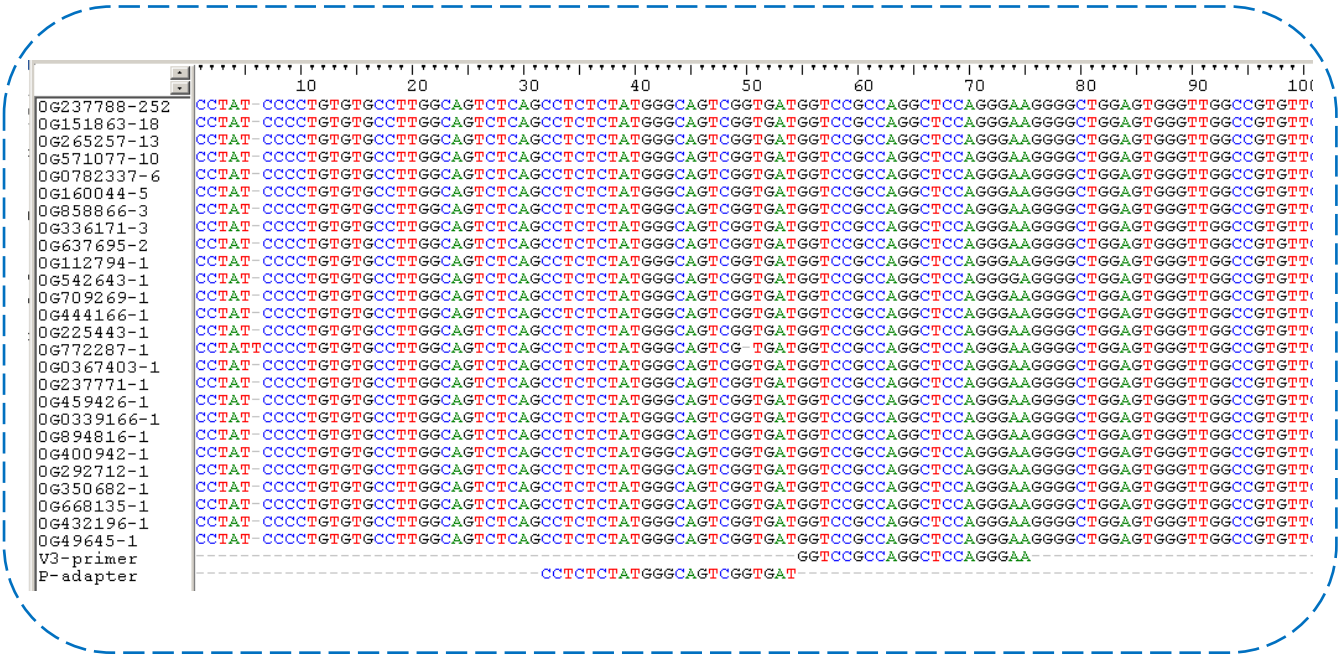


Figure 42: Closer analysis of individual sequences from highly mutated IgM peak observed in Tick2 indicates that this exceptionally high mutation load is not biological. *This effect is probably due to the fact that conensus sequences have not been trimmed after the V-primer sequence. Alignments of next generation sequencing reads at the V-region side with the V3-region primer and P1 Ion Torrent adapter are depicted.*

4.8 Closer analysis of largely distorted IgM repertoires of two donors with a recent tick bite

When analyzing samples with our multicolor flow cytometry panel (Figure 13), one donor with a recent tick bite strongly deviated from the others. Perturbations were mainly observed in non-switched B cell subsets but were very constant over the three timepoints. Furthermore, the participants generally recognized and removed the tick early, which was only attached between 0-3 days. For these reasons, we initially thought that this rather special B cell subset composition is specific for that donor and has nothing to do with the tick bite. When we observed a largely skewed B cell repertoire

in two other donors with a recent tick bite with our next generation sequencing approach however, we thought it might be worth further investigating this phenomenon. In accordance with the multicolor flow cytometry data (Figure 13C), that indicated a very high percentage of naïve B cells in the peripheral blood of this donor (84% versus an average of 63% for all samples), the final IgM libraries from this donor were very concentrated (Figure 43E) and the raw reads contained a high percentage of singlet or doublet UIDs (Figure 43D) but also gave a high number of consensus sequences (Figure 43B). The two donors with a skewed IgM repertoire (from which we unfortunately don't have FACS data) on the other hand showed only a very low percentage of reads occurring as singlet or doublet UIDs (Figure 43D), plus their final IgM libraries were not very concentrated (Figure 43E). The final concentrations of the IgG libraries were however similar to other samples (Figure 43F), indicating that it was not a general problem with the sample (like integrity of the RNA). In case of very low input material, one could imagine that errors occurring during PCR amplification might artificially induce clonal expansions. In order to exclude this, we had a look at the UID distributions of the samples from the sequencing Chips of the tick bite donors and saw that they were not particularly different from the other samples that were run on the same Chip (Figure 44A). In fact, the UID distributions of the largely expanded clones nicely reflected the distribution of the rest of the sample (Figure 44B). Also the UIDs of these individual sequences were very different from each other (Figure 44C), indicating that the final consensus reads really originated from different mRNA molecules. The mutation status of these sequences indicated that they are quite distant from the germline, while this was generally not true for related sequences found in other donors (Figure 45), indicating that these mutations are probably true mutations and were not solely present due to difficulties of the Ion Torrent to sequence reads of that particular composition. In accordance with the high number of sequences belonging to isotype mixed clusters in these two donors with a recent tick bite (Figure 37A), in both we found (although at very low levels) IgG sequences related to these large clones and which showed the same distance from the germline as their IgM counterparts. In the expanded clone from Tick2, replacement mutations were mainly concentrated in the CDR2 region (Figure 46), which would make biologically sense. We focused here on the CDR2 as the large difference between CDR3 regions and their original germline D genes generally makes it difficult to assess mutation rates in this region. In Tick1 some amino acid changes in the CDR2 were also found, replacement mutations were however also scattered along the whole sequence (Figure 46).

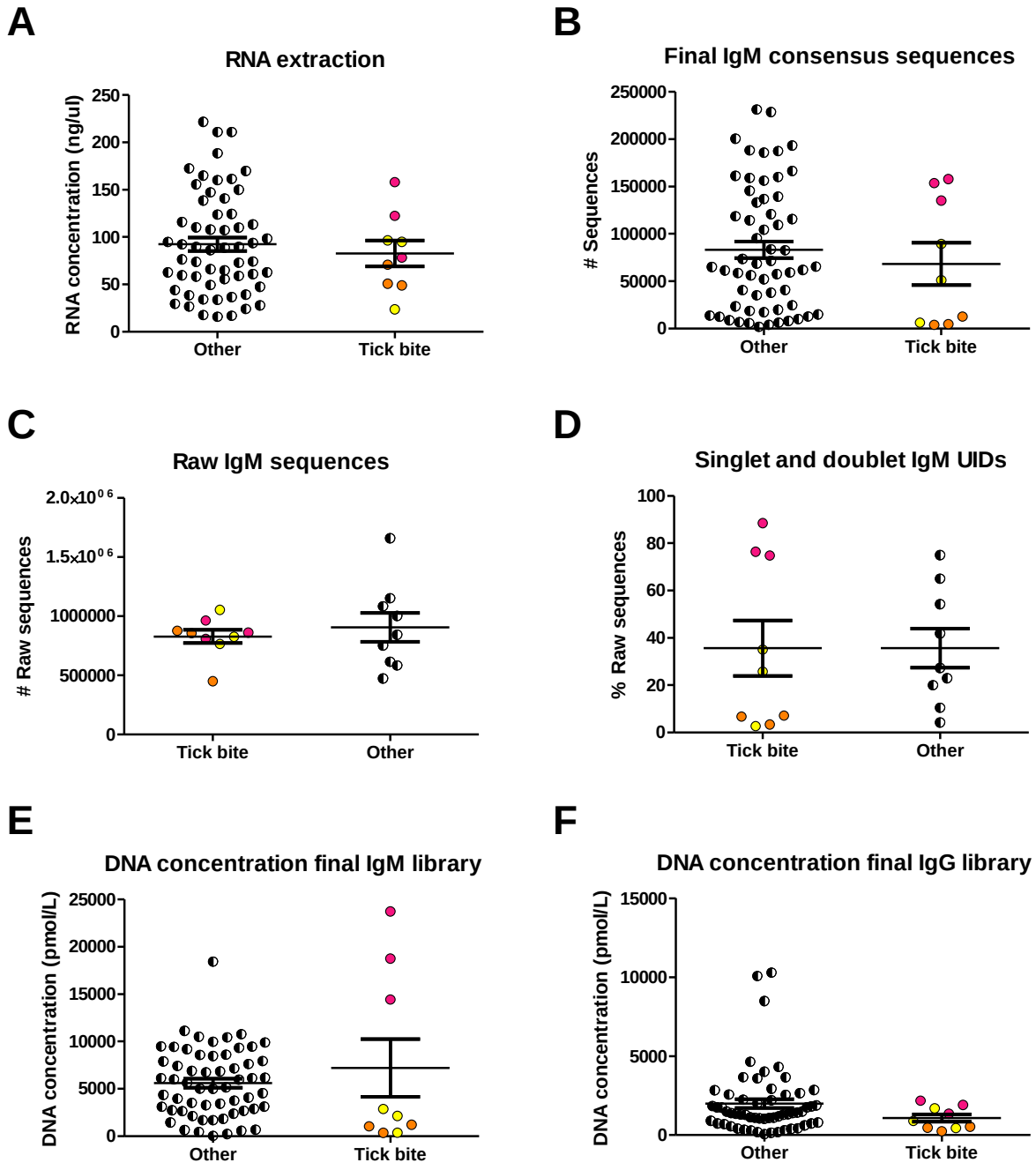


Figure 43: Exclusion of sample preparation or sequencing biases as reason for outlying tick bite samples. *Comparison of RNA extraction (A), final consensus sequences (B), raw sequences (C), percentage reads that showed up as singlet or doublet UIDs (D) and DNA concentration of final libraries (E) between IgM samples of donors with a recent tick bite and either all the other samples (A, B, E, F) or samples sequenced in the same sequencing run (C,D). (F) DNA concentration of final IgG libraries. As the aim was to see whether the tick bite samples are outliers as compared to the rest of the samples, no statistical test was performed. Yellow: Tick1. Orange: Tick2. Pink: Tick3.*

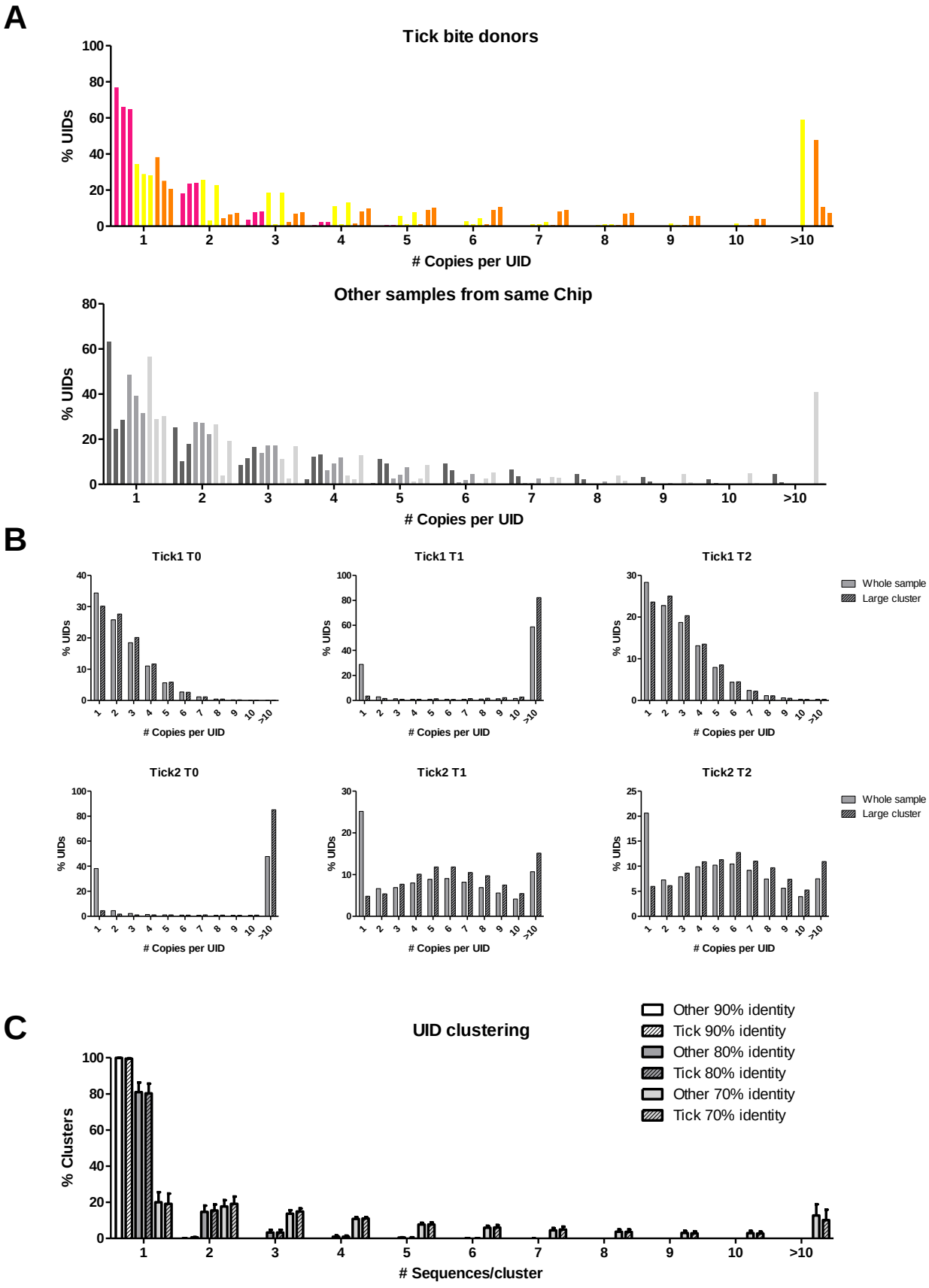

A
Tick bite donors
% UIDs
# Copies per UID
Other samples from same Chip
% UIDs
# Copies per UID
B
Tick1 T0
Tick1 T1
Tick1 T2
Tick2 T0
Tick2 T1
Tick2 T2
% UIDs
# Copies per UID
Whole sample
Large cluster
C
UID clustering
% Clusters
# Sequences/cluster
Other 90% identity
Tick 90% identity
Other 80% identity
Tick 80% identity
Other 70% identity
Tick 70% identity

Figure 44: Exclusion that low input material might have generated artificial clones due to PCR errors in UID barcodes. *(A) Percentage of UIDs occurring at the indicated copy numbers for samples from tick bite donors (top) and corresponding acute patient samples that were sequenced on the same Chip (bottom). (B) Comparison of UID distribution from large clusters detected in Tick1 and Tick2 with those from the whole sample. (C) Cd-hit [336,337] was used to cluster UIDs from sequences of the large clusters from Tick1 and Tick2 into groups using different degrees of similarity. To test whether the UID distributions deviate from those of other samples, the same number of UIDs was randomly selected (using shuf command in bash) from the other samples sequenced on the same Chip (Other) and clustering compared to that one of the tick bite samples (Tick).*

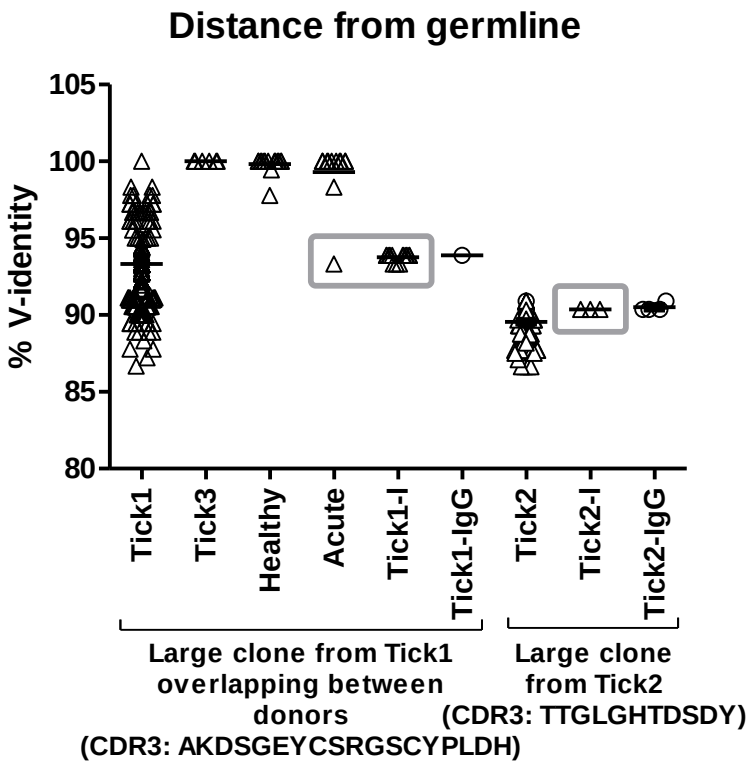


Figure 45: Mutation levels of large clusters identified in the IgM repertoires of Tick1 and Tick2. *Mutation levels in sequences related [12,127,171] (Figure 35) to the large clusters were assessed by determining V-identities using IMGT/High V-QUEST [354]. For each sample, sequences were manually cleaved off with the BioEdit Sequence Alignment Editor until the V-region primer sequence and collapsed before determining V-identity distributions. Triangles: IgM isotype. Circles: IgG isotype. Tick1-l and Tick2-l: only considering sequences present at very high copy numbers. Sequences related to the largely expanded clone from Tick1 were also found in Tick3 as well as in acute and healthy donor samples. V-identities of these related sequences are also represented (Tick3, Healthy, Acute). Replacement mutations depicted in Figure 46 were determined from the sequences highlighed in grey boxes.*

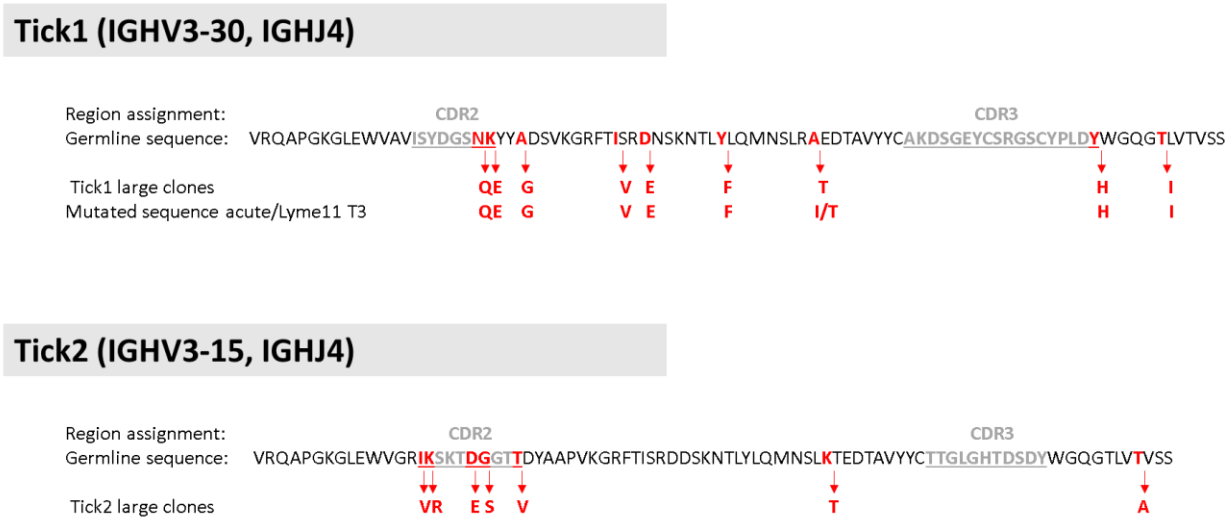


Figure 46: Analysis of replacement mutations occurring in the clones from Tick1 and Tick2 that predominate their IgM repertoires. *Replacement mutations (highlighted in red) were determined manually by comparison of the sequences present at high copy numbers to the germline at the amino acid level using BioEdit Sequence Alignment Editor. Lyme11 T3: a long term follow up sample from Lyme 11 (donor with a CLL clone) was available and a related sequence was found in that sample.*

4.9 Determination of VlsE-C6 memory B cell levels in deep sequencing data

We tested whether the peptide tetramer reactive single cell CDR3 sequences can be retrieved in the deep sequencing data. Again, the vast majority (81%, 130/160) of the single cell sequences that could be found in the NGS data clustered with sequences of repertoires from the same donor (Peptide positive same donor: 96, Control same donor: 34) while only a few clustered with sequences from another donor (Peptide positive different donor: 15, Control different donor: 15). In acute donors, sequences belonging to peptide positive clusters did not exceed 1.5% of the total (meaning pooled IgM and IgG data) repertoire (Figure 47A). Clusters belonging to peptide positive sequences from different individuals made up less than 0.1% of the pooled repertoire samples (Figure 47A). This further indicates that the repertoire is very “private”. The clusters related to sequences from the peptide enriched B cells were generally not among the top clusters of the acute patients. Although more sequences from the bulk B cell repertoire clustered together with peptide associated sequences compared to control sequences, this difference may be partially explained by differences in sequence numbers. When we randomly select the same number of peptide and control sequences per timepoint, the observed effect is gone. Since in that case the number of sequences left to analyze is very low, a higher number of cells would be needed to correctly address this issue. When we are looking at the

average cluster frequency of individual clusters however, clones related to peptide reactive B cells have a tendency to make up a larger proportion of repertoire sequences as compared to randomly sorted control cells (Figure 47B), indicating some degree of selection. A slightly larger fraction of peptide positive single cell sequences clustered with deep sequencing data as compared to randomly sorted cells (Figure 47C). When excluding all the single cell sequences related to the CLL clone, only 4% of the peptide positive single cells were found to cluster among each other and could also be found in the NGS data (Figure 47D). Phylogenetic tree analysis was performed on clones for which a large number of related sequences was found in the NGS data (Figure 48). Generally clones were composed either of IgM or of IgG isotypes and correlated well with the isotypes of the single cells. Only in two of the six donors analyzed in that way could we find isotype mixed clusters (Lyme3 and Lyme8). As for the two largely expanded clones from the tick bite individuals (Figure 45), the clones were predominated by IgM sequences and contained only very few IgG counterparts. In two donors we found only IgM clusters (Lyme4 and Lyme 7). In the last third of these patients we were able to find both IgM and IgG clusters (Lyme6 and Lyme9).

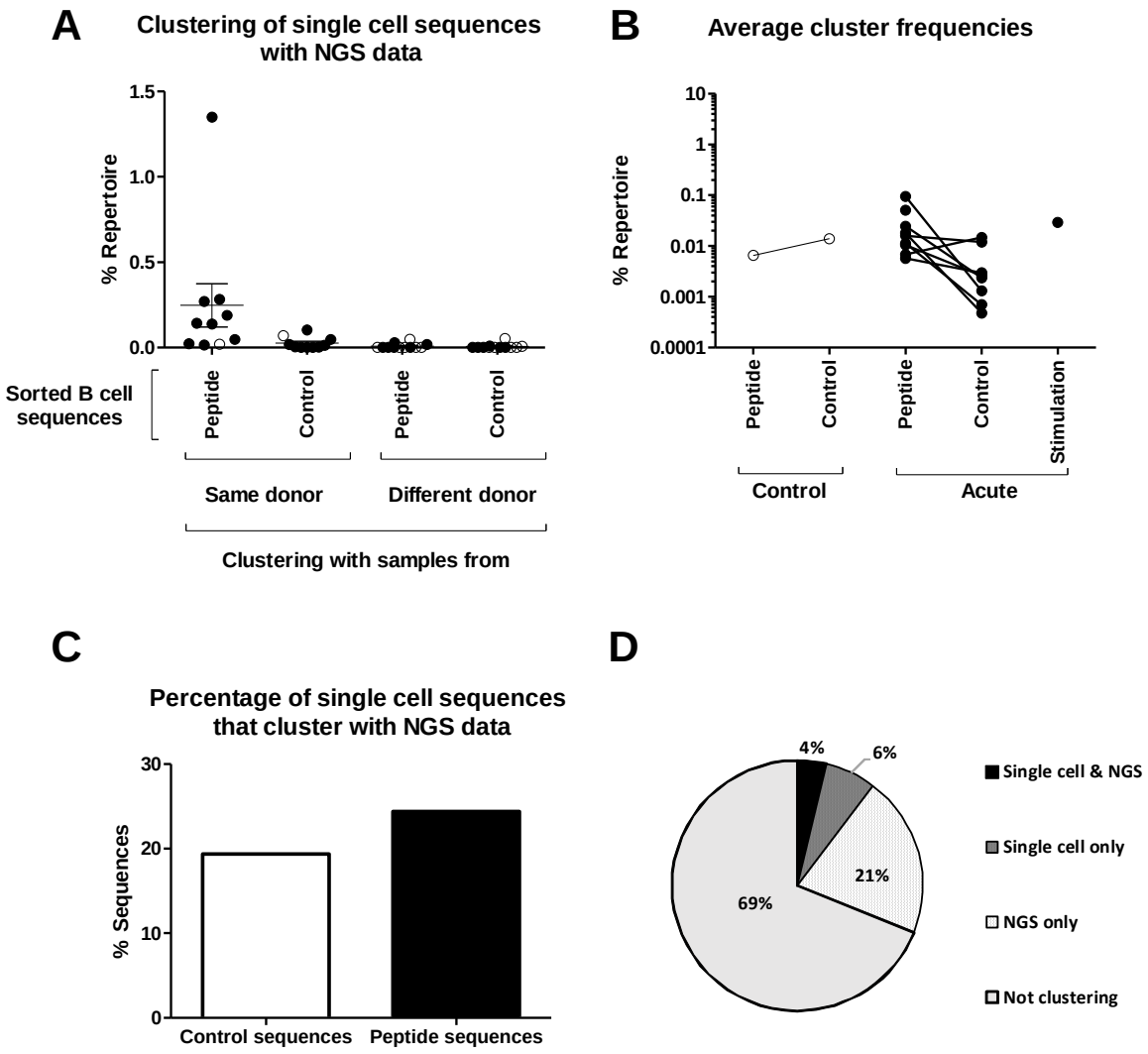
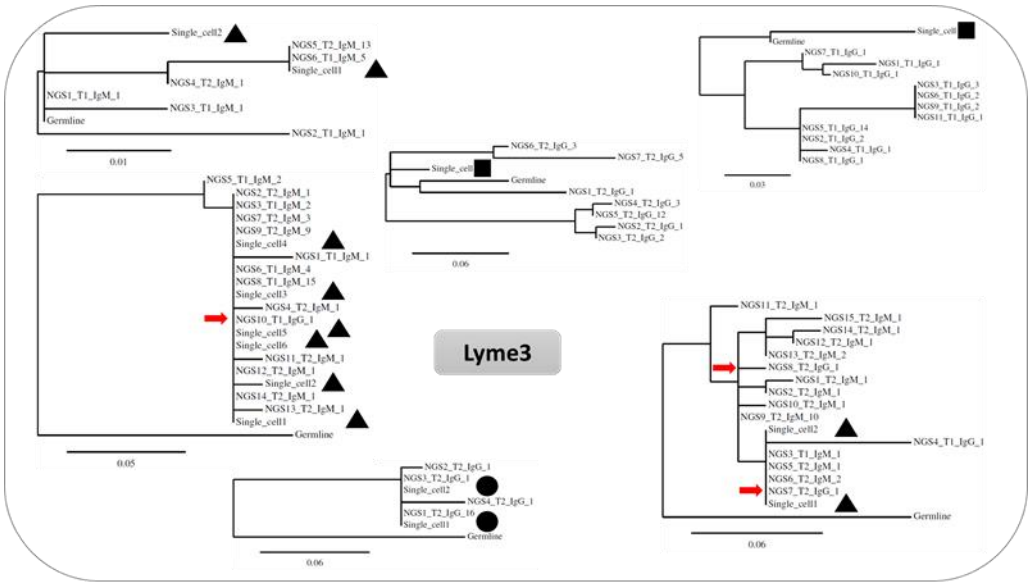
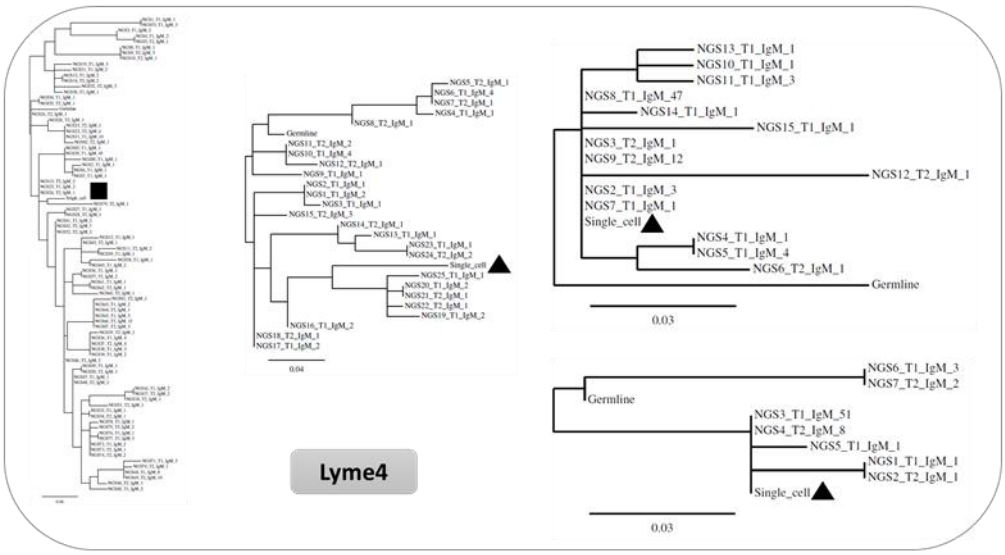
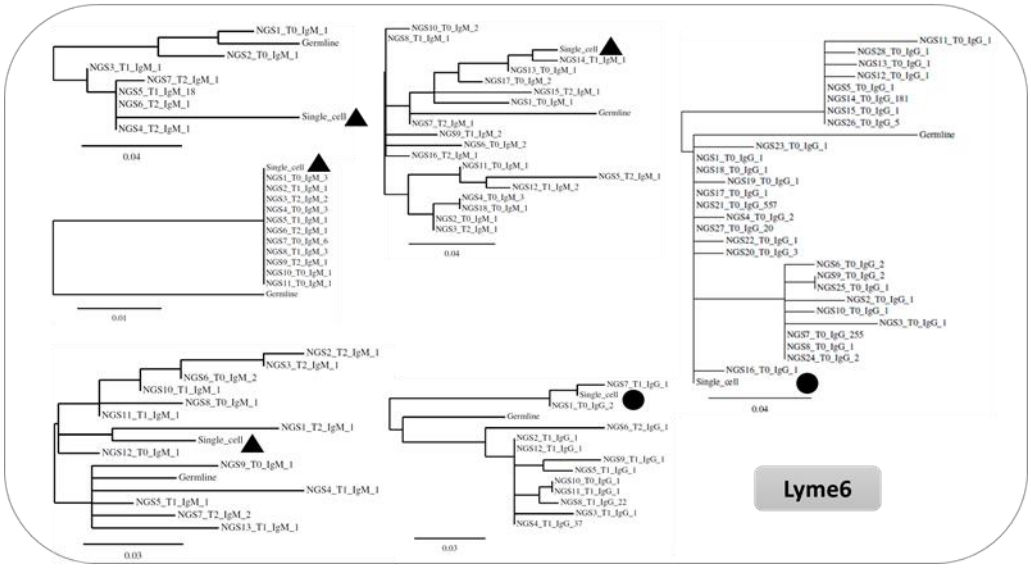


Figure 47: Clustering of single cell sequences with deep sequencing data. (*A) Change-o [338] was used to cluster pooled NGS sequences (referred here to as "repertoire", including pooled IgM and IgG sequences from all sequenced timepoints) from each donor with those originating from the isolated single cells. The percentage repertoire that is related to the isolated single cells is represented. Values were separated according to whether the sequences were found back in the same donor from which they were isolated or in a different donor. Open circles: Control samples. Black circles: acute patient samples. (B) Average frequencies of individual single cell clusters found in the NGS data from the same donor were determined. Values obtained for randomly sorted control (Control) and VlsE-C6 epitope (Peptide) positive cells are depicted separately. Values obtained from the same donors are connected by lines. The cluster size from one VlsE-C6 reactive sequence isolated with in vitro single cell stimulation (please refer to section 4.10 Confirmation of findings with in vitro single cell stimulation) that was found back in the deep sequencing data is also represented (Stimulation). (C) Percentage of single cell sequences that cluster with deep sequencing data. (D) Percentage of peptide positive single cell sequences clustering with next generation sequencing data and/or among each other.*


Lyme3


Lyme4


Lyme6

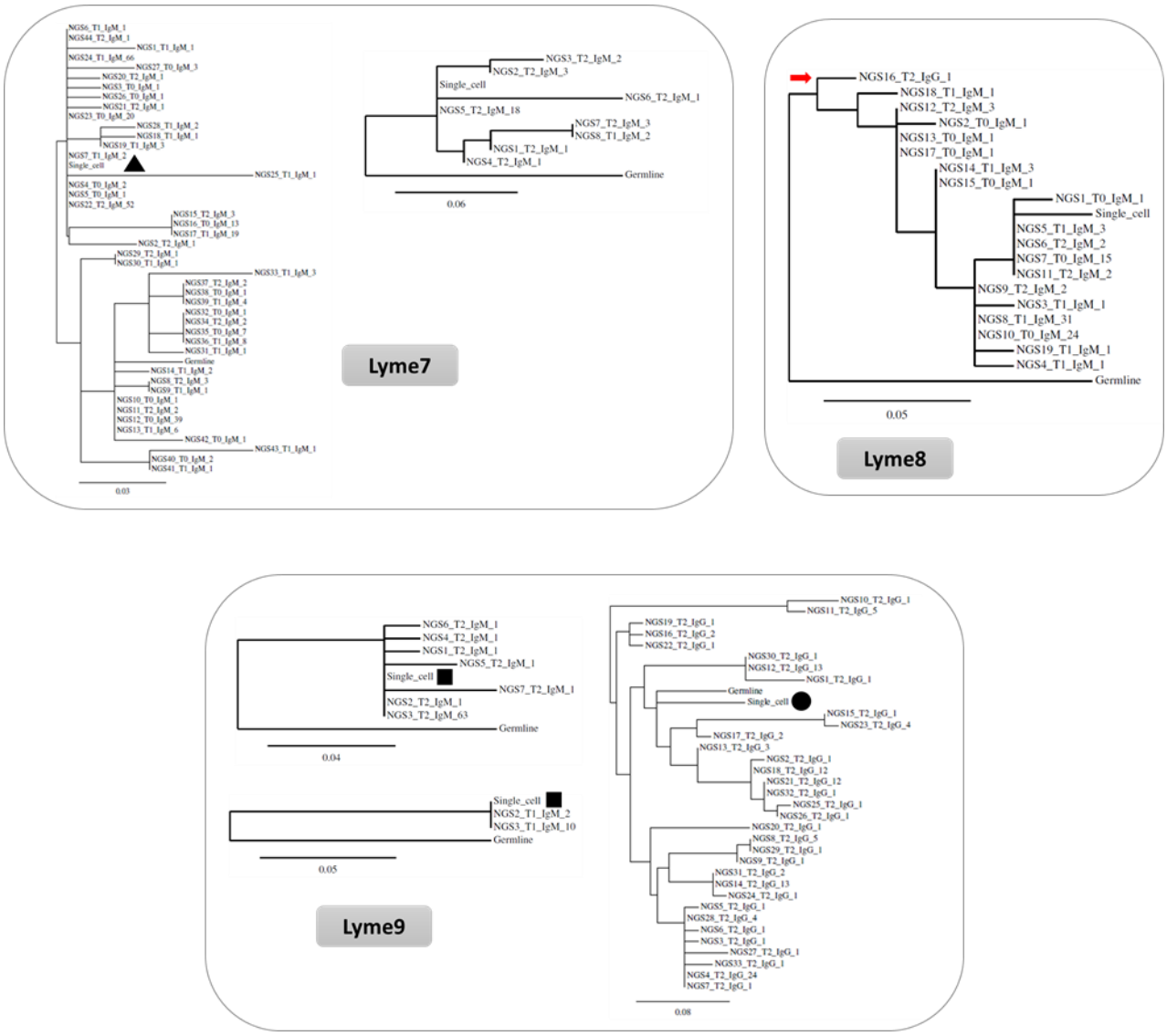


Figure 48: Phylogenetic tree analysis of clones releated to VlsE-C6-tetramer positive B cells. *The isotypes of the single cells are indicated either as black circle (IgG), triangle (IgM) or square (IgA). IgG sequences found in IgM clusters are highlighted with red arrows.*

4.10 Confirmation of findings with *in vitro* single cell stimulation

In order to validate the isolated VlsE-C6 reactive B cell receptor sequences with a second method, IgG^+ B cells were sorted by negative selection and single cells stimulated for antibody production in 384 well plates. The *in vitro* stimulation experiments were carried out by Alessia Colone. On average 10 plates (~3800 cells) per patient have been seeded. Lyme8 and Lyme4 were selected for this experiment, because they showed changes in antibody titers over time (Table 2), indicating an ongoing immune response. Lyme3 and Lyme6 have been selected, because they were among the

patients in which we found most single cell sequences back in the IgG deep sequencing data (Figure 48), indicating that they might have expanded clones of interest in their repertoire. In addition to that, Lyme3 also showed a large cluster of IgM peptide sorted B cells.

Supernatants of stimulated B cells from these four donors were screened for reactivity against the three *Borrelia* VlsE-C6 epitopes and the B cell receptor of cells from positive wells were sequenced. To confirm specificity of these cells for *Borrelia*, the supernatants were also screened for a whole panel of other antigens to which patients very likely have been exposed to in the past (Figure 49). All the wells except one were confirmed to be negative for the other tested antigens. Only one well reacted with both *Borrelia* and measles virus (CDR3-IMGT: ARADDSPSYYVNAFDL). With this approach, we were able to extract another 8 sequences from VlsE-C6 reactive B cells, with one of them occuring in two different wells from the same donor (Table 7). When comparing the frequency of VlsE-C6 reactive B cells to that of cells reacting with antigens from past infections or vaccinations (Figure 49A), the percentage of VlsE-C6 reactive wells was similar or even lower, confirming the low frequency of these cells among PBMCs. On average 0.07% of wells reacted with VlsE-C6 (Lyme8: 0.142% (IP90,PT7,B31); Lyme4: 0.071% (IP90), 0.047% (PT7,B31); Lyme6: 0.032% (IP90,PT7), 0.095% (B31); Lyme3: 0.028% (IP90,B31), 0.057% (PT7)). *Borrelia* reactive B cells harboured a similar amount of mutations in their immunoglobulin heavy chain sequences as compared to those that reacted with the other antigens tested (Figure 49B). When analyzing their sequences, we were able to confirm the heterogeneous nature of VlsE-C6 reactive B cell clones (Table 7). Only one of these sequences - the same sequence that was present in two different VlsE-C6-reactive wells – was related (1 mismacthes within CDR3 was allowed) to a cluster (0.0292% repertoire) from the deep sequencing data of the same donor (Figure 50), further confirming that sequences from *Borrelia* reactive B cells are patient specific and generally not present among the most abundant B cell receptor mRNAs of the repertoire.

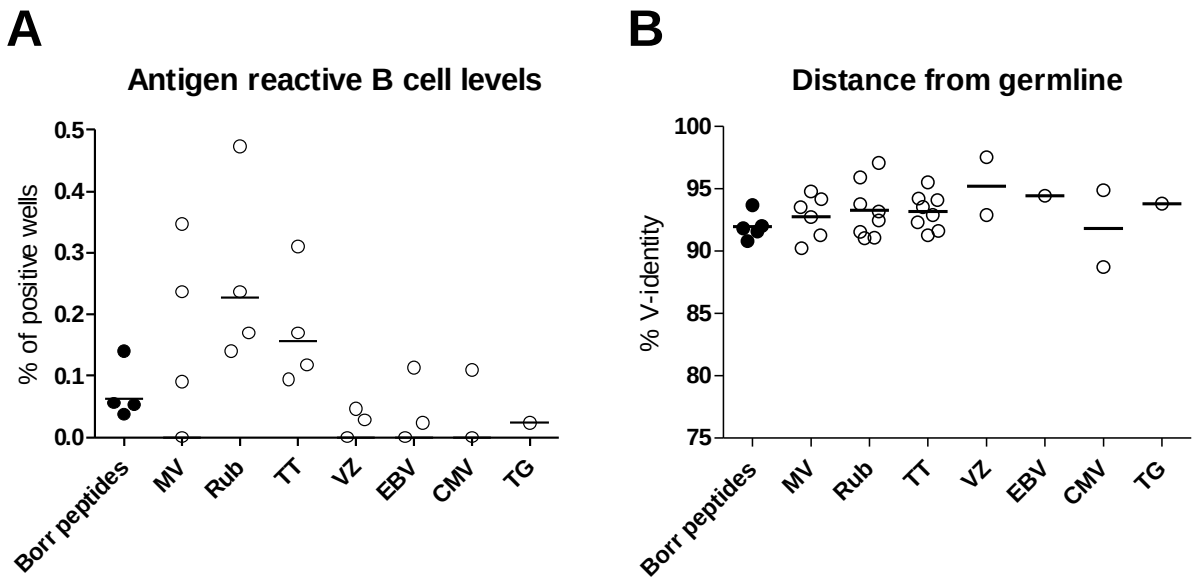


Figure 49: Frequency of antigen specific single B cells and mutation levels in their immunoglobulin heavy chain sequences. *(A) Percentage wells from single cell stimulation that were positive in ELISAs for the indicated antigens or pathogens. Borr peptides: Borrelia VlsE-C6 peptides from either B31, PT7 or IP90 strain [280]; MV: measles virus; Rub: Rubella; TT: Tetanus toxoid; VZ: Varicella-Zooster virus; EBV: Epstein Barr virus; CMV: Cytomegalovirus; TG: Toxoplasma Gondii. Filled black circles: Borrelia peptides; Open circles: other antigens or pathogens. These results were kindly provided by Alessia Colone. (B) V-identity of the single cell sequences obtained as determined by IMGT/HighV-QUEST [341].*

Table 7: VlsE-C6 reactive immunoglobulin heavy chain sequences isolated with *in vitro* single cell stimulation.

CDR3	Stimulated sample	# Sequences	Related clone NGS	V-gene	J-gene
ARVVCTGGT	Lyme8 (T1)	1	no	IGHV4-34	IGHJ5
ARGTRDGQ	Lyme8 (T1)	1	no	IGHV3-33	IGHJ1
ARGTRDGQNPEFDY	Lyme8 (T1)	2	Lyme8	IGHV3-33	IGHJ4
ARADDSPSYYVNAFDL	Lyme4 (T1)	1	no	IGHV5-51	IGHJ3
AILWGGLGVVTPDYYYFEMDV	Lyme3 (T1 & T2)	1	no	IGHV1-69	IGHJ6
TRNEIQLDD	Lyme6 (T0 & T1 & T2)	1	no	IGHV3-49	IGHJ4
VRNHTVYVGGNCNSNPGVFDI	Lyme6 (T0 & T1 & T2)	1	no	IGHV3-48	IGHJ3
ARSYEYYYYYAMDV	Lyme6 (T0 & T1 & T2)	1	Control (1 sequence)	IGHV3-7	IGHJ6

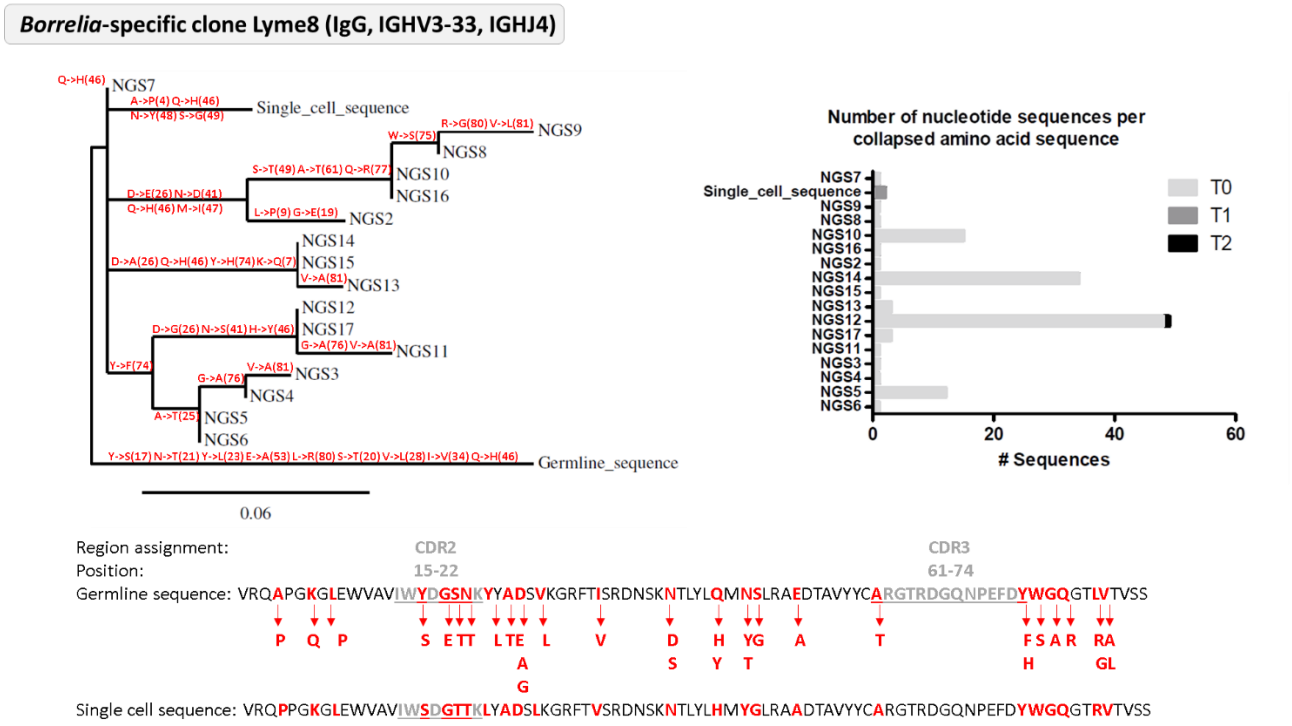


Figure 50: Closer analysis of *Borrelia* VlsE-C6 reactive B cell clone found back in the deep sequencing data from Lyme 8. *Top left: Phylogenetic tree highlighting the mutations (with concerned positions in brackets) occurring at each branch of the tree. Top right: Number of nucleotide sequences that made up each amino acid sequence represented in the phylogenetic tree as well as the timepoint at which they were present. Sequences have the same order than in the phylogenetic tree. Bottom: VDJ sequence of the germline and the isolated single cell sequences with amino acid substitutions from that clone highlighted in red.*

4.11 Clustering of isolated IgG$^+$ memory B cell receptor sequences of different specificities with next generation sequencing data

In order to get a first impression on the level of the CDR3s isolated with *in vitro* single cell stimulation in our next generation sequencing data we used a simple agrep command allowing one mismatch to screen for similar CDR3 amino acid sequences. In accordance with what we found with this approach for VlsE-C6 reactive B cells (Table 7, Figure 50), when searching for clones with CDR3 sequences related to single cells reactive towards the other antigens, we also hardly found any sequences back in the NGS data. These results confirm, that our deep sequencing approach is focusing on expanded or activated clones but is missing a large portion of the memory B cell repertoire. In one patient (Lyme 4), we however observed an expanded TT reactive B cell clone. This single cell sequence has been isolated from that same donor. In order to get additional evidence for a possible expansion of

TT reactive B cell clones in Lyme 4, we searched our NGS data for published CDR3 sequences from TT reactive B cells [326] and also found higher levels in the dataset from that same acute donor as compared to the rest of the samples. In order to generate evidence for an ongoing TT specific immune response, we analyzed the mutation status and isotypes of the extracted sequences (Figure 51). Although TT related NGS sequences could be found in the majority of our samples, they were rather composed of unmutated sequences of IgM isotype as would be expected by naïve B cells potentially able to start an immune response towards TT. Interestingly, two related sequences of IgG isotype could be extracted, which showed a mutation level similar to other isotype switched cells. Although one expanded clone related to published CDR3s could be identified in Lyme 4, this clone was of IgM isotype and only slightly mutated. The clone that was related to the isotype switched B cell isolated with *in vitro* stimulation on the other hand showed a higher mutation level that was similar to the sequence from the single cell and also the *Borrelia* reactive clone extracted from Lyme 8 (which showed the highest mutation level of all the clones).

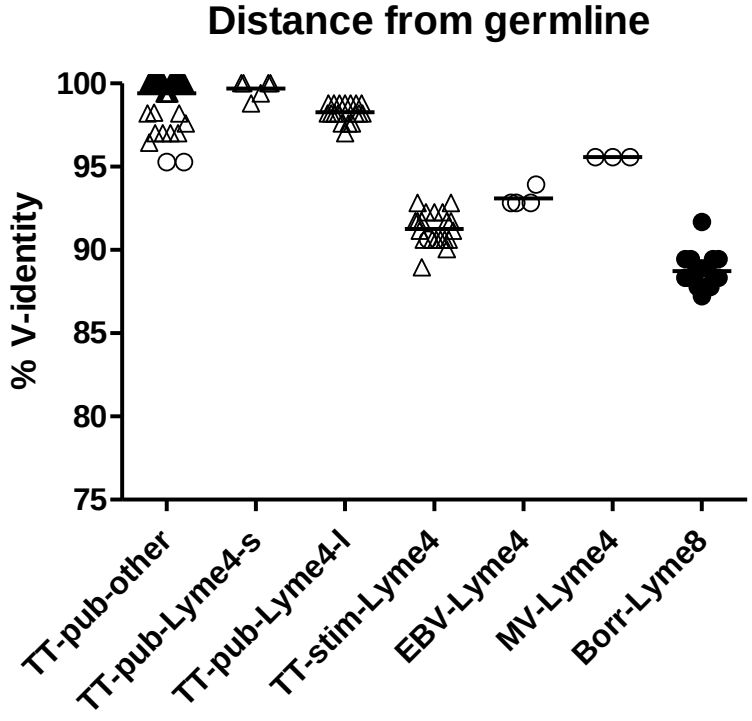


Figure 51: V-identities and isotypes of sequences from next generation sequencing data that are related to isolated single cell sequences. *TT-pub-other: TT-related sequences found in other donors than Lyme 4 using the published CDR3 dataset [326], TT-pub-Lyme 4-s: small TT related clones found back in the deep sequencing data from Lyme 4 using the published CDR3 dataset; TT-pub-Lyme 4-l: large TT related clone found back in the deep sequencing data from Lyme 4 using the*

published CDR3 dataset; TT-stim-Lyme4: Clone related to the CDR3 sequence isolated in this study from Lyme4 by in vitro single cell stimulation; EBV-Lyme4 and MV-Lyme4: clones related to the single cells reactive with EBV and MV isolated by in vitro single cell stimulation from Lyme 4. Borr-Lyme8: Clone related to Borrelia VlsE-C6 reactive single cell sequence found in Lyme 8. Triangles: IgM isotype, Circles: IgG isotype, Filled circles: Borrelia peptide reactive IgG clone, Open circles or triangles: clones related to other antigens or pathogens.

Phylogenetic tree analysis was performed on the clones for which we could find related sequences in the next generation sequencing data (Figure 52). The phylogenetic tree related to the stimulated cell shows sequences going into different directions from the germline, which is indicative of an ongoing immune response (Figure 52A). A closer mutation analysis confirmed, that the majority of mutations within the TT related clone were occurring in the CDR2. This region was very different between the isolated single cell sequence and the assigned germline V gene (Figure 52B). This confirms, that the observed mutations might indeed be biologically significant and not due to sequencing errors.

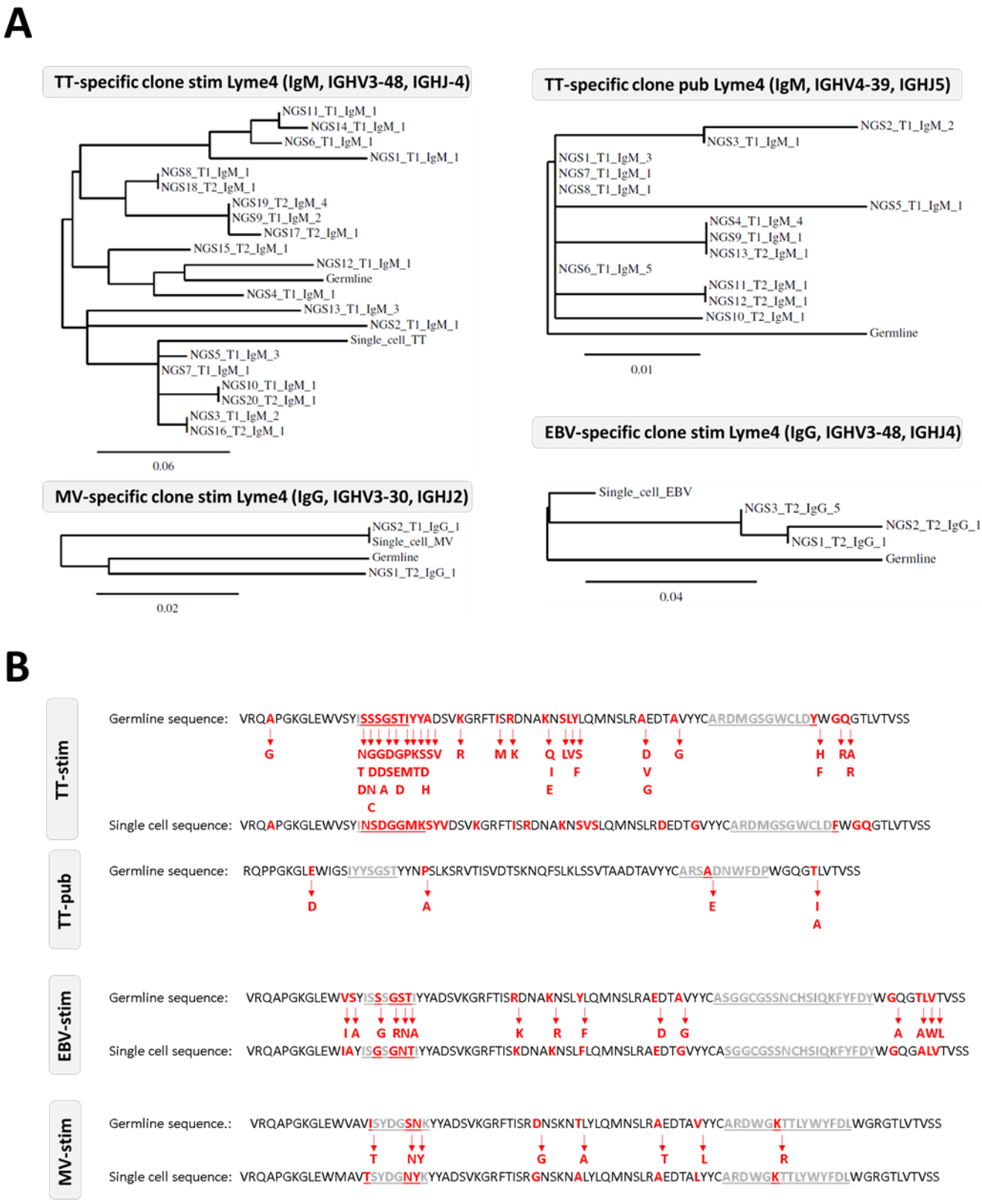
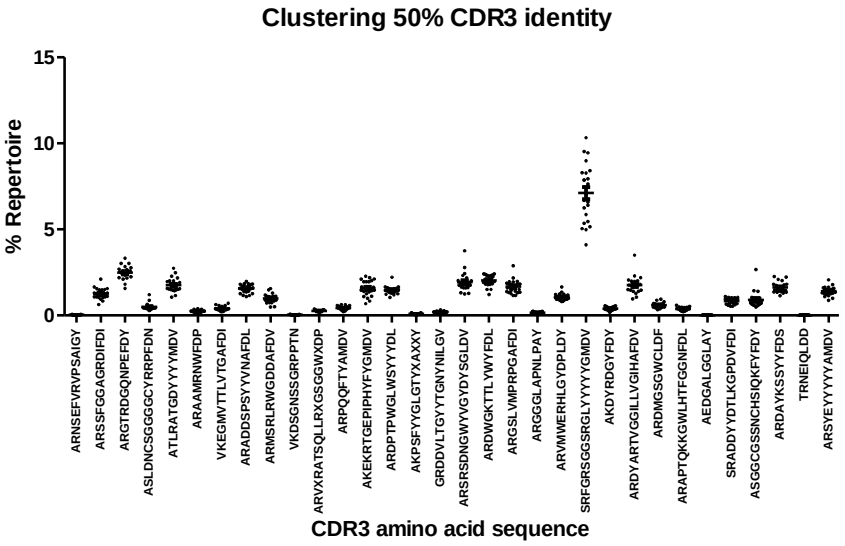


Figure 52: Closer analysis of clones from deep sequencing data related to B cells that are reactive against antigens other than the *Borrelia* VlsE-C6 epitope. *(A) Phylogenetic trees showing the relations between individual sequences of the selected clones. The timepoint, isotype and number*

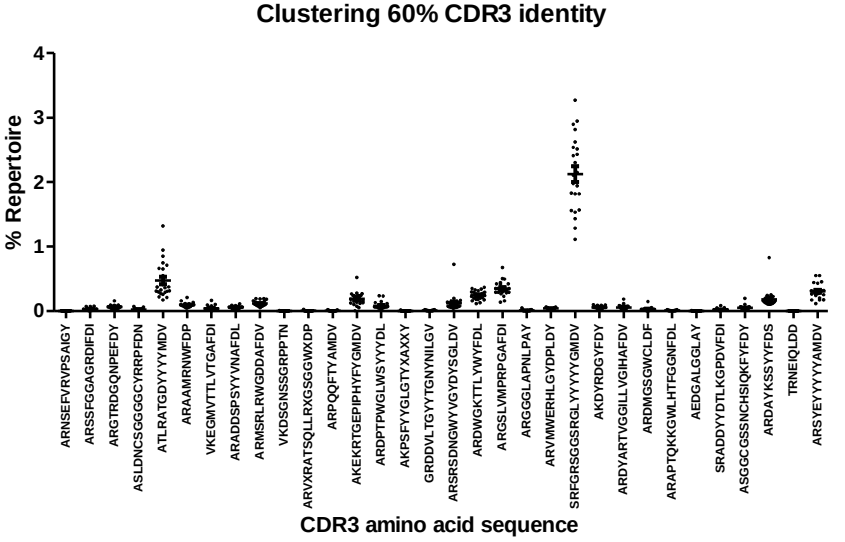
of nucleotide sequences are indicated in the name of each amino acid sequence from the phylogenetic trees. (B) VDJ germline and single cell sequences with amino acid substitutions highlighted in red.

Since this was a rather rough approach to search for related CDR3 sequences, we used Cd-hit-2d [336,337] to cluster CDR3s according to different CDR3 amino acid identity thresholds (Figure 53). At lower thresholds, the extent of clustering with deep sequencing data was different between CDR3s but for each individual CDR3, the level was similar among the different donors (Figure 53A-C). At these levels of identities, CDR3s containing stretches of Tyrosines seem to cluster generally with more sequences from each repertoire than other CDR3s. At the highest CDR3 identity thresholds (Figure 53E-F), the CDR3s already found with the simpler agrep command were sticking out also when analyzing with this method (*Borrelia* CDR3: ARGTRDGQNPEFDY; MV CDR3: ARDWGKTTLYWYFDL; EBV CDR3: ASGGCGSSNCHSIQKFYFDY; TT CDR3: ARDMGSGWCLDF).

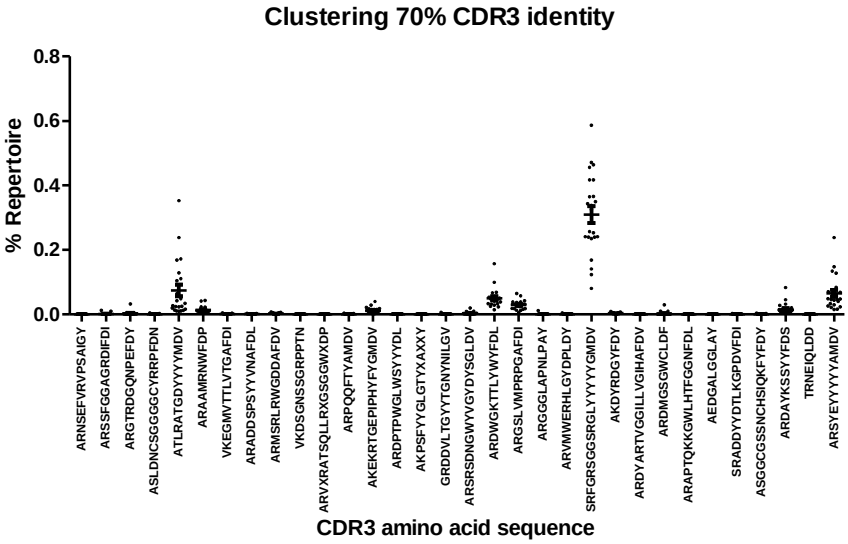
A


Clustering 50% CDR3 identity
% Repertoire
15
10
5
0
ARNSEFVRVPSAIGY
ARSSFGGAGRDIFDI
ARGTRDGQNPEFDY
ASLDNCSGGGGCYRRPFDN
ATLRATGDYYYYMDV
ARAAMRNWFDP
VKEGMVTTLVTGAFDI
ARADDSPSYYVNAFDL
ARMSRLRWGDDAFDV
VKDSGNSSGRPPTN
ARVXRATSQLLRXGSGGWXDP
ARPQQFTYAMDV
AKEKRTGEPIPHYFYGMDV
ARDPTPWGLWSYYYDL
AKPSFYYGLGTYXAXXY
GRDDVLTGYYTGNYNILGV
ARSRSDNGWYVGYDYSGLDV
ARDWGKTTLYWYFDL
ARGSLVMPRPGAFDI
ARGGGLAPNLPAY
ARVMWERHLGYDPLDY
SRFGRSGGSRGLYYYYYGMDV
AKDYRDGYFDY
ARDYARTVGGILLVGIHAFDV
ARDMGSGWCLDF
ARAPTQKKGWLHTFGGNFDL
AEDGALGGLAY
SRADDYYDTLKGPDVFDI
ASGGCGSSNCHSIQKFYFDY
ARDAYKSSYYFDS
TRNEIQLDD
ARSYEYYYYYAMDV
CDR3 amino acid sequence

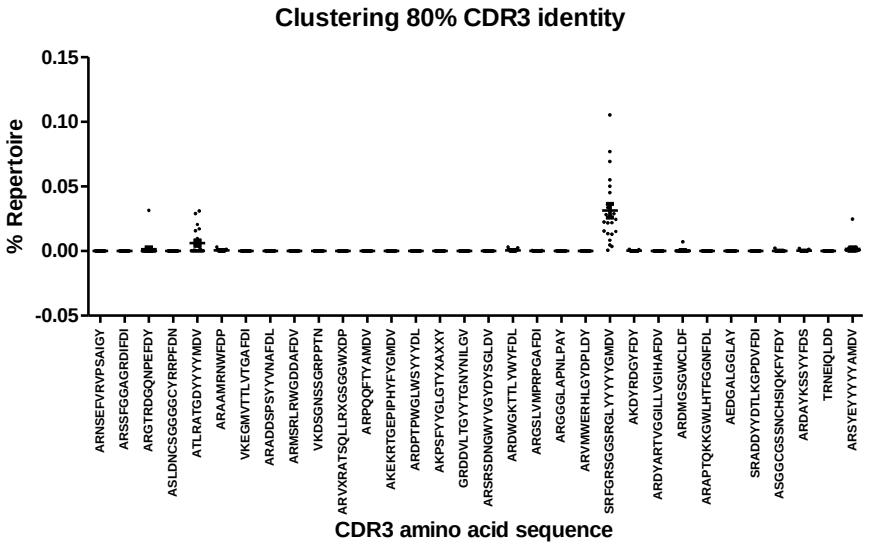
B


Clustering 60% CDR3 identity
% Repertoire
4
3
2
1
0
ARNSEFVRVPSAIGY
ARSSFGGAGRDIFDI
ARGTRDGQNPEFDY
ASLDNCSGGGGCYRRPFDN
ATLRATGDYYYYMDV
ARAAMRNWFDP
VKEGMVTTLVTGAFDI
ARADDSPSYYVNAFDL
ARMSRLRWGDDAFDV
VKDSGNSSGRPPTN
ARVXRATSQLLRXGSGGWXDP
ARPQQFTYAMDV
AKEKRTGEPIPHYFYGMDV
ARDPTPWGLWSYYYDL
AKPSFYYGLGTYXAXXY
GRDDVLTGYYTGNYNILGV
ARSRSDNGWYVGYDYSGLDV
ARDWGKTTLYWYFDL
ARGSLVMPRPGAFDI
ARGGGLAPNLPAY
ARVMWERHLGYDPLDY
SRFGRSGGSRGLYYYYYGMDV
AKDYRDGYFDY
ARDYARTVGGILLVGIHAFDV
ARDMGSGWCLDF
ARAPTQKKGWLHTFGGNFDL
AEDGALGGLAY
SRADDYYDTLKGPDVFDI
ASGGCGSSNCHSIQKFYFDY
ARDAYKSSYYFDS
TRNEIQLDD
ARSYEYYYYYAMDV
CDR3 amino acid sequence

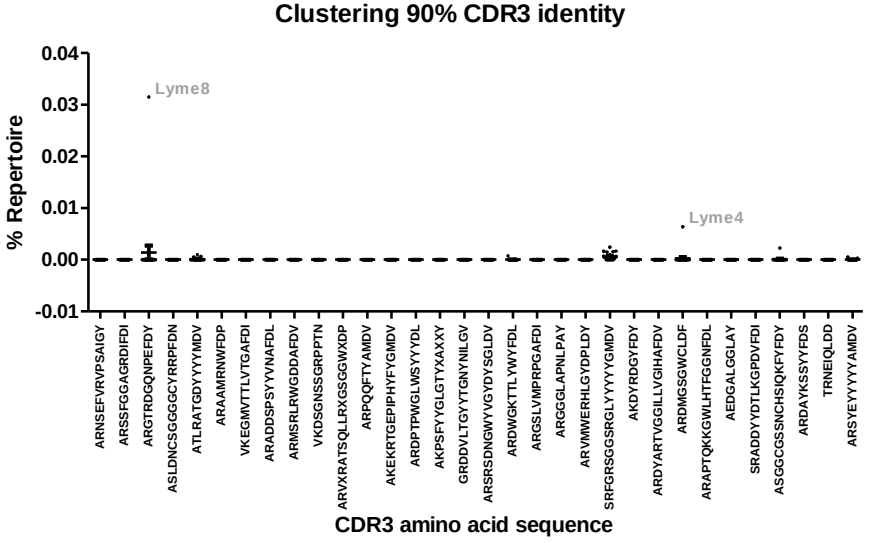
C


Clustering 70% CDR3 identity
% Repertoire
0.8
0.6
0.4
0.2
0.0
ARNSEFVRVPSAIGY
ARSSFGGAGRDIFDI
ARGTRDGQNPEFDY
ASLDNCSGGGGCYRRPFDN
ATLRATGDYYYYMDV
ARAAMRNWFDP
VKEGMVTTLVTGAFDI
ARADDSPSYYVNAFDL
ARMSRLRWGDDAFDV
VKDSGNSSGRPPTN
ARVXRATSQLLRXGSGGWXDP
ARPQQFTYAMDV
AKEKRTGEPIPHYFYGMDV
ARDPTPWGLWSYYYDL
AKPSFYYGLGTYXAXXY
GRDDVLTGYYTGNYNILGV
ARSRSDNGWYVGYDYSGLDV
ARDWGKTTLYWYFDL
ARGSLVMPRPGAFDI
ARGGGLAPNLPAY
ARVMWERHLGYDPLDY
SRFGRSGGSRGLYYYYYGMDV
AKDYRDGYFDY
ARDYARTVGGILLVGIHAFDV
ARDMGSGWCLDF
ARAPTQKKGWLHTFGGNFDL
AEDGALGGLAY
SRADDYYDTLKGPDVFDI
ASGGCGSSNCHSIQKFYFDY
ARDAYKSSYYFDS
TRNEIQLDD
ARSYEYYYYYAMDV
CDR3 amino acid sequence

D


Clustering 80% CDR3 identity
% Repertoire
0.15
0.10
0.05
0.00
-0.05
ARNSEFVRVPSAIGY
ARSSFGGAGRDIFDI
ARGTRDGQNPEFDY
ASLDNCSGGGGCYRRPFDN
ATLRATGDYYYYMDV
ARAAMRNWFDP
VKEGMVTTLVTGAFDI
ARADDSPSYYVNAFDL
ARMSRLRWGDDAFDV
VKDSGNSSGRPPTN
ARVXRATSQLLRXGSGGWXDP
ARPQQFTYAMDV
AKEKRTGEPIPHYFYGMDV
ARDPTPWGLWSYYYDL
AKPSFYYGLGTYXAXXY
GRDDVLTGYYTGNYNILGV
ARSRSDNGWYVGYDYSGLDV
ARDWGKTTLYWYFDL
ARGSLVMPRPGAFDI
ARGGGLAPNLPAY
ARVMWERHLGYDPLDY
SRFGRSGGSRGLYYYYYGMDV
AKDYRDGYFDY
ARDYARTVGGILLVGIHAFDV
ARDMGSGWCLDF
ARAPTQKKGWLHTFGGNFDL
AEDGALGGLAY
SRADDYYDTLKGPDVFDI
ASGGCGSSNCHSIQKFYFDY
ARDAYKSSYYFDS
TRNEIQLDD
ARSYEYYYYYAMDV
CDR3 amino acid sequence

E


Clustering 90% CDR3 identity
% Repertoire
0.04
0.03
0.02
0.01
0.00
-0.01
Lyme8
Lyme4
ARNSEFVRVPSAIGY
ARSSFGGAGRDIFDI
ARGTRDGQNPEFDY
ASLDNCSGGGGCYRRPFDN
ATLRATGDYYYYMDV
ARAAMRNWFDP
VKEGMVTTLVTGAFDI
ARADDSPSYYVNAFDL
ARMSRLRWGDDAFDV
VKDSGNSSGRPPTN
ARVXRATSQLLRXGSGGWXDP
ARPQQFTYAMDV
AKEKRTGEPIPHYFYGMDV
ARDPTPWGLWSYYYDL
AKPSFYYGLGTYXAXXY
GRDDVLTGYYTGNYNILGV
ARSRSDNGWYVGYDYSGLDV
ARDWGKTTLYWYFDL
ARGSLVMPRPGAFDI
ARGGGLAPNLPAY
ARVMWERHLGYDPLDY
SRFGRSGGSRGLYYYYYGMDV
AKDYRDGYFDY
ARDYARTVGGILLVGIHAFDV
ARDMGSGWCLDF
ARAPTQKKGWLHTFGGNFDL
AEDGALGGLAY
SRADDYYDTLKGPDVFDI
ASGGCGSSNCHSIQKFYFDY
ARDAYKSSYYFDS
TRNEIQLDD
ARSYEYYYYYAMDV
CDR3 amino acid sequence

F

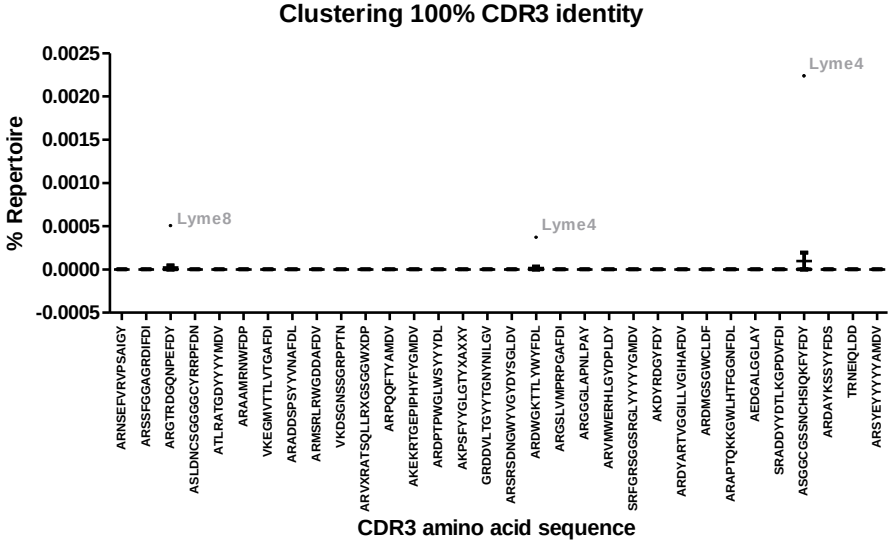

Clustering 100% CDR3 identity
% Repertoire
0.0025
0.0020
0.0015
0.0010
0.0005
0.0000
-0.0005
Lyme8
Lyme4
Lyme4
ARNSEFVRVPSAIGY
ARSSFGGAGRDIFDI
ARGTRDGQNPEFDY
ASLDNCSGGGGCYRRPFDN
ATLRATGDYYYYMDV
ARAAMRNWFDP
VKEGMVTTLVTGAFDI
ARADDSPSYYVNAFDL
ARMSRLRWGDDAFDV
VKDSGNSSGRPPTN
ARVXRATSQLLRXGSGGWXDP
ARPQQFTYAMDV
AKEKRTGEPIPHYFYGMDV
ARDPTPWGLWSYYYDL
AKPSFYYGLGTYXAXXY
GRDDVLTGYYTGNYNILGV
ARSRSDNGWYVGYDYSGLDV
ARDWGKTTLYWYFDL
ARGSLVMPRPGAFDI
ARGGGLAPNLPAY
ARVMWERHLGYDPLDY
SRFGRSGGSRGLYYYYYGMDV
AKDYRDGYFDY
ARDYARTVGGILLVGIHAFDV
ARDMGSGWCLDF
ARAPTQKKGWLHTFGGNFDL
AEDGALGGLAY
SRADDYYDTLKGPDVFDI
ASGGCGSSNCHSIQKFYFDY
ARDAYKSSYYFDS
TRNEIQLDD
ARSYEYYYYYAMDV
CDR3 amino acid sequence

Figure 53: Clustering of single cell CDR3 amino acid sequences with next generation sequencing data using Cd-hit-2d [336,337]. *Different CDR3 amino acid identity thresholds were used to cluster single cell sequences with next generation sequencing data (pooled IgM and IgG samples). For every CDR3 analyzed, each dot corresponds to the clustering value of one donor.*

Since at lower CDR3 amino acid identity thresholds we observed a similar increase in clustering of sequences from the repertoires of the different donors, it was difficult to determine a clustering threshold that would allow to separate truly related sequences from randomly clustering ones. Only for the CDR3s reactive with *Borrelia* peptides could we try to set a threshold by testing whether there is a difference in clustering between the two groups at a certain value. At the ideal threshold, we would expect a maximum difference between the two groups. When doing this analysis with three different *Borrelia* related CDR3s (Figure 54), we observed no clear difference in clustering at the lower identity thresholds between the groups, indicating that the majority of these sequences are probably not truly related to each other. Individual acute Lyme disease donors stood out at the different thresholds, but no general difference between the two groups could be found. At higher identity thresholds, only the CDR3 that clustered before with the NGS data from Lyme8 showed a clear difference in Lyme8 as compared to all the other samples. No increased clustering was observed with any of the other CDR3s or identity thresholds tested, indicating that these CDR3s have been rather uniquely produced in the donor from which they were isolated and that sequences that were not found back at higher identity thresholds can also not be detected as common to the acute Lyme disease group when allowing a greater amount of mismatches.

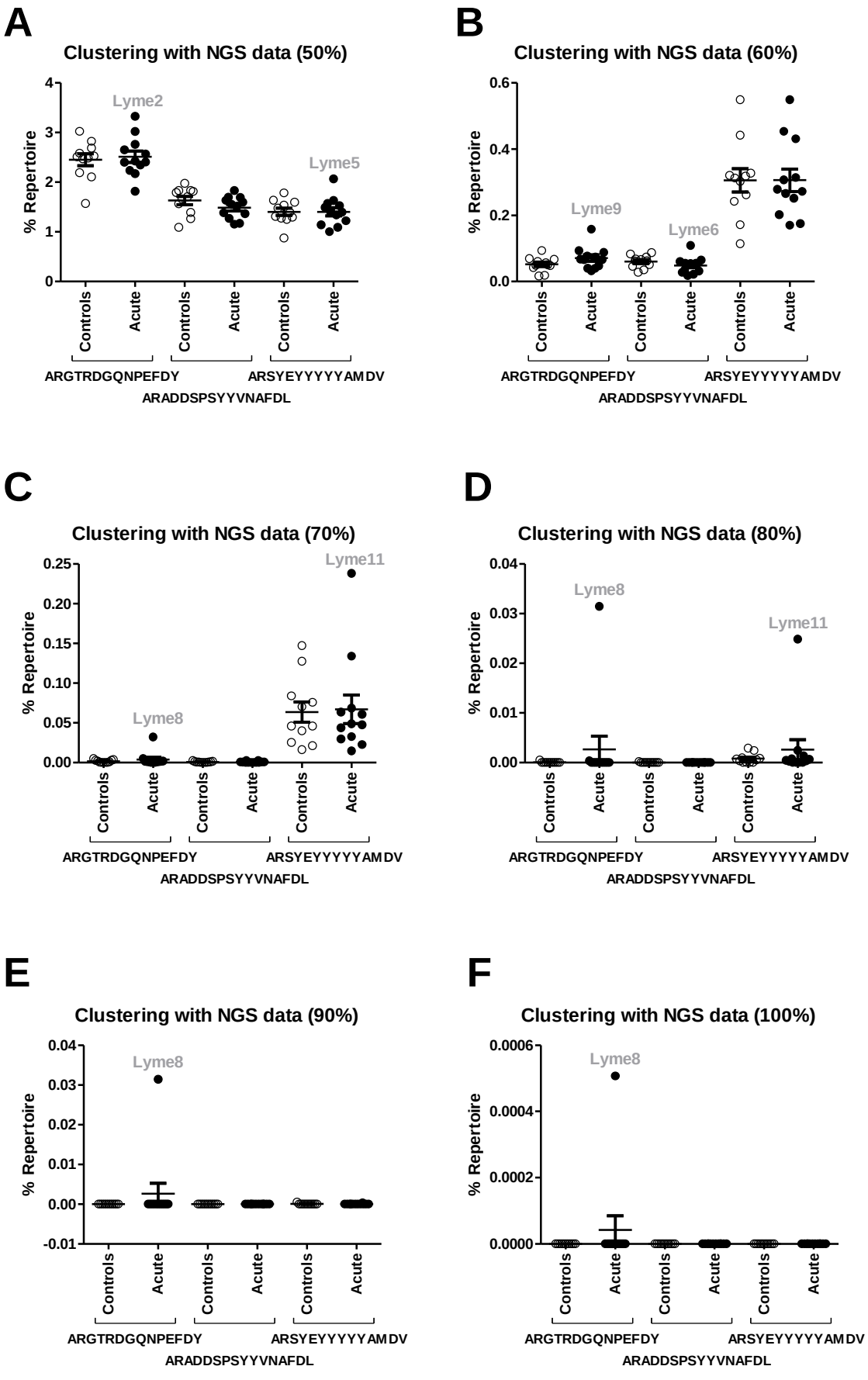

A
Clustering with NGS data (50%)
% Repertoire
Lyme2
Lyme5
Controls
Acute
ARGTRDGQNPEFDY
ARSYEYYYYYAMDV
ARADDSPSYYVNAFDL
B
Clustering with NGS data (60%)
% Repertoire
Lyme9
Lyme6
C
Clustering with NGS data (70%)
% Repertoire
Lyme11
Lyme8
D
Clustering with NGS data (80%)
% Repertoire
Lyme8
Lyme11
E
Clustering with NGS data (90%)
% Repertoire
Lyme8
F
Clustering with NGS data (100%)
% Repertoire
Lyme8

Figure 54: Comparison of clustering of CDR3 amino acid sequences from VlsE-C6 reactive single B cells with next generation sequencing data between acute and control subjects. *Different CDR3 identity thresholds were tested using Cd-hit-2d [336,337].*

5 DISCUSSION

5.1 A heterogeneous B cell immune response associated with low numbers of antigen-reactive B cells of interest in peripheral blood complicates the extraction of Lyme disease specific B cell repertoire signatures

Our data indicate, that the human B cell immune response to *Borrelia* might be very heterogeneous when comparing different donors. By analyzing peripheral blood B cell subsets with multicolor flow cytometry, we hoped to identify a B cell subset that would be uniquely expanded in the acute Lyme disease condition but not in control samples. This would have allowed us to focus on a B cell subset of interest thereby avoiding sequencing of B cell receptors that are not of direct interest. A subset of the analyzed acute Lyme disease patients showed slightly elevated levels of plasmablasts, CD23 expressing naïve B cells and IgG positive memory B cells, which are indicative of an ongoing B cell immune response. The rather unique B cell subset compositions of the individual donors made it however difficult to reliably extract one subpopulation that is clearly expanded upon *Borrelia* infection.

The absence of detectable distortions in B cell subpopulations could be explained by a combination of the low number of patients analyzed and either a general low level of B cell activation [251] or heterogeneous responses between patients [304]. Considering the possibility that a robust B cell activation occurs in only a subset of Lyme disease patients is legitimate, as previous studies on human antibody responses classified patients into three different categories of responders [304]. Our analysis enabled us to get some insights into the peripheral blood plasmablast cell numbers to be expected in our samples. Using conventional gating strategies based on CD27 and CD38 as markers, their numbers were generally very low (between 0-0.64% of B cells). This is in accordance with previous studies that showed an unstable cell surface marker expression on these cells after cryopreservation [363]. An extensive amount of gating and back-gating needed to be done in order to find this cell subpopulation and separation was not totally clear, indicating that it might be difficult to sort [334]

this cell subset from our patient samples. Since only surface marker expression seems to be affected, frozen PBMCs can still be used to study B cell repertoire perturbations induced by immune stimuli [1].

When comparing the IgM heavy chain B cell repertoires of acute Lyme disease patients to healthy controls, we identified an increased CDR3 amino acid overlap between different timepoints. Interestingly, healthy individuals showed the highest interdonor IgM CDR3 amino acid overlap. This would be in line with the hypothesis, that the expanded clones in acute and tick bite individuals are rather patient unique. In healthy individuals we expect more naïve B cells, which are less mutated and might therefore have a higher chance to overlap between individuals. Generally interdonor overlap was much higher for IgM as compared to IgG samples. Furthermore, closer analysis of individual IgM clones indicated, that related naïve sequences can be found in the repertoires of other donors. For a cluster to be included into the group of overlapping sequences it is sufficient that only one related naïve B cell sequence is present in the repertoire of the other donor. This might explain why we don't see any difference between the groups. Sorting of CD27+ memory B cells will be required to correctly address this issue.

In accordance with a slight increase of IgG expressing memory B cells, we detected an increased diversity and interdonor cluster overlap of acute patients' IgG repertoires. Also when assessing CDR3 amino acid overlap between samples from different timepoints, we observed an increased overlap between acute patients' samples as compared to the controls. Final IgG libraries prepared from acute patient samples also showed a tendency towards a higher DNA concentration, further indicating expanded IgG expressing B cell clones. Since the observed differences were hardly detectable and diversity seemed to rather have increased, we came to the conclusion, that *Borrelia* induced clones might not necessarily predominate the repertoires.

In the present study, we did not normalize for the same number of reads per sample. For several reasons, we concluded that it might be important to include all the obtained reads into the analysis. First, 86% (32/37) of the sequences from single memory B cells isolated by *in vitro* single cell stimulation in this study could not be found back in the next generation sequencing data. From this, we concluded that we are far from covering the whole memory B cell repertoire. Recent studies however showed that B cell clones of low frequency are more likely to overlap between donors [171].

As the same library preparation protocol (starting from PBMCs from the same amount of blood and using the same RNA input concentration) was used for all the samples, differences in final concentrations of libraries are likely due to real biological differences. Correct normalization is important for a reliable interpretation of the results, but is rather difficult due to the complex nature of the B cell repertoire. The repertoire is composed of a mixture of B cells, some of which could still be expanded from previous infections or be induced by other ongoing immune responses. Different donors also might react with different magnitudes or kinetics to the same stimulus [4,169]. Clones induced by other infections can largely influence the results of both normalized but also non-normalized samples. Imagine, that there is a largely expanded clone form another infection still in the repertoire, this influences the percentage and rank of a *Borrelia* reactive clone. Even the number of expressed *Borrelia* proteins could influence these parameters. When working with whole PBMC samples, the presence of other immune cells and the relative abundances of IgM and IgG expressing B cells are also influencing the results. Since at this time the specificity of the expanded clones cannot be directly inferred from the deep sequencing data, it remains to be investigated, whether the clones that are predominating the individual repertoires have been induced by *Borrelia* or by some other immune stimulus. The most expanded clusters of each IgG repertoire did not show any overlap with other acute Lyme disease patient samples. When focusing on the top 50 clusters of each sample 58 of 1700 (34 samples; 3.4%) were overlapping with sequences from another acute donor. Corresponding sequences from the other repertoires were however generally not belonging to the top 50 clusters. When doing this analysis, two clusters particularly stood out. The first interesting cluster (CDR3aa: X-K(R)-W-R-X-X-Q-S-E-X-D(E)-X; IGHV3-7; IGHJ4) was found in 5 different donors (Top50: Lyme7(T2), Lyme12(T1); Top200: Lyme4(T2), Lyme11(T0); low level: Lyme1(T0,T1,T2), Lyme4(T1), Lyme12(T0,T2), Lyme7(T0)) and the second one (CDR3aa: A-K-X-X-X-X-X-C-S-X-X-X-C-Y-X-F-D-X, IGHV3-23; IGHJ4) was found to overlap even between 6 different donors (Top50: Lyme3 (T2), Lyme9 (T2); Top200: Lyme9(T0), Lyme 8 (T0), Lyme12(T0); low level: Lyme2(T1); Lyme7(T0,T1)). Both these clusters could not be found in any of the control samples. These clusters have been isolated by a rather rough manual approach and more sophisticated statistical analysis taking the occurrences and levels of individual clusters as well as general clonal repertoire distributions into account will be required to provide more reliable results. Isolation of antigen reactive single B cells and confirmation of their B cell receptor specificity is a more direct method that is crucial to solve these issues. Unfortunately we did not have the time to perform confirmatory expression experiments anymore, but when isolating VlsE-C6 reactive B cells we found that they are very heterogeneous between donors. Undersampling of peptide reactive B cells with our

single cell approaches is one possible explanation for the observed absence of clustering among peptide reactive sequences. Since next generation sequencing allows to analyze a much larger proportion of the repertoire, we tested to what extent we could find sequences related to the peptide enriched B cells in our deep sequencing data. Also with this approach, the vast majority of the single cell sequences clustered with sequences of repertoires from the same donor (86%, 96 of the 111 that were found back in the deep sequencing data). Overall our data indicate, that the B cell repertoire is rather "personal", showing little overlap between individuals. These findings do however not exclude the possibility that *Borrelia* reactive B cell repertoire signatures might exist. Our data just indicate, that these signatures are more complex than initially imagined and that one might need to dig deeper into the repertoires in order to identify them. Specificity of extracted candidate signatures needs to be confirmed by different experiments. Quantitative PCR approaches allowing the amplification of the individual clones of interest from samples might be more sensitive than sequencing of bulk B cell repertoires.

Since there cannot be an indefinite number of possibilities to bind to the same epitope, it is probably just a matter of how many sequences and patients will be analyzed before finding clones overlapping between individuals that were challenged with the same antigen. This has in fact been nicely demonstrated by recent next generation sequencing studies, which were able to show that elevated levels of antigen related sequences indeed occur after corresponding immune stimuli [169,171,326]. Although a rather clear immune signature could be isolated from a large acute Dengue patient cohort based on NGS data alone [165], the specificity of this sequence for this disease still needs to be confirmed, as it was proposed that the Dengue virus activates natural IgG B cells with polyreactive properties [166]. It would make sense, that evolution has designed these types of antibodies in a way that they share more common features between donors as compared to other more specific antibodies. The possibility of polyreactive sequences being included into vaccine induced clusters has already been proposed previously and should be considered as possible part of any antigen specific B cell immune response [126]. Interestingly very old mouse studies already indicated, that thousands of different B cell clones might be capable of reacting with an antigen and that B cell immune responses should rather be seen as being composed of a mixture of antibodies with antigen- and poly-specific reactivities [100]. Given, that such a mechanism would provide a broader range of protection, it would only make biologically sense to induce an immune response that is at the same time specific but that can also confer protection against related (possibly mutated versions of the same) pathogens. Although repertoires of donors challenged with the same immune stimuli seem to have a higher

chance to overlap, it remains open whether this degree of overlap is sufficient and specific enough for repertoire sequences to be useful to support diagnosis or to say something about the immunological history of patients. Plasma cells isolated after meningococcal vaccination for example only shared 0.17-2.2% CDR3 sequences between donors, indicating that the vast majority of generated clones are rather patient unique [127]. Furthermore, when analyzing the presence of hepatitis B virus reactive antibody sequences in immunized individuals, a large fraction was already present in the switched repertoire before the immune challenge, showing the existence of cross-reactive sequences in the memory repertoire of individuals [126]. In the present study, we also found that sequences overlapping between a higher number of donors were generally found in both groups (those with acute Lyme disease but also in the control group), raising the question of whether clustering approaches and overlap analysis in combination with such low numbers of patients are rather enriching for polyreactive B cell repertoire sequences. Overlapping sequences between acute patients and controls could be explained in that case by a subclinical infection or still expanded clones from previous infections in the repertoires of healthy donors [12]. Our multicolor flow cytometry data indicate, that indeed a subset of healthy donors might have expanded clones and an ongoing B cell immune response. This might be another explanation for why we hardly detected any difference between acute Lyme disease patients and healthy individuals when assessing repertoire diversity. Similar to other studies, our results are indicating, that cluster overlap is rather occurring at the lower level memory B cell subpopulation [171]. Although a closer analysis of the individual overlapping clusters is required to completely confirm this statement, our data showed rather an anticorrelation between CDR3 amino acid overlap among different timepoints as compared to overlap between different donors. As antigen-specific memory B cells can remain in the repertoire for even a lifetime [144], it remains to be investigated, whether the expansion of these cells upon acute infection is large enough for them to be useful for diagnosis. Furthermore - coming back to the problem with polyspecificity - since the major role of memory B cells is to elicit a quicker response to reinfection, which can occur by related but not identical pathogens, they have been shown to be less specific for the antigen as compared to the antibody secreting cells which have the main role to confer protection against the same pathogen [56,68-70]. On the other hand polyreactive memory B cell clones capable of recognizing different mutant variants of the same pathogen might be advantageous for diagnosis, because only one clone would need to be identified for an acute infection caused by the different variants. The kinetics of memory B cells is rather slow and thereby maybe not suitable for diagnosis [116]. Given all these considerations, the challenges of future studies will be to isolate those rare expanded B cell clones that overlap between donors, but which at the same time show a high degree

of specificity for the antigen. From this and previous studies, we can only conclude that the number of possible ways of our B cell repertoire to recognize an antigen is probably very large. Mathematical modelling and deeper statistical analysis will be required to successfully address this issue but is out of the scope of the present study. To reliably determine the sizes of B cell repertoires reactive towards different antigens or epitopes, single cell isolation techniques will need to be combined with next generation sequencing approaches [123,124] allowing to screen larger numbers of cells and to perform robust statistical analysis.

Our deep sequencing data suggest that the switched memory and plasma cells generated in response to *Borrelia* infection might be less mutated than those generated in response to other infections or vaccinations. Low mutation frequencies - and hence low specificity - could be one explanation for the high level of crossreactivity of acute Lyme disease patient sera [285], which renders the development of specific serological tests difficult. On the other hand, direct comparison to other primary immune responses is crucial to be able to make such a statement. Since the majority of the patients form our study encountered Lyme disease for the first time, it is also possible that the low mutation frequency is due to the nature of a primary immune response rather than problems to undergo efficient germinal center reactions. As these sequences coincide with the major IgG peak and are more mutated than the ones from the IgM repertoire, they have probably undergone some degree of selection in germinal centers or similar structures. For VlsE-C6 reactive memory B cells we observed a similar or even increased mutation level as compared to memory B cells of other specificities, indicating that at least B cell clones reacting towards this epitope harbour “normal” mutation levels.

In accordance with previous studies [3,126] which already indicated, that the effect of immune stimuli on the B cell repertoire is expected to be small and not easily detectable, we were not able to observe drastic changes in repertoire diversities upon *Borrelia* infection. The number of final sequences can influence diversity indices and differences observed between biological replicates were in a similar range than differences observed between samples, indicating that the effect of *Borrelia* on the B cell repertoire might be smaller than inherent fluctuations of the data. Also when performing hierarchical clustering using diversity profiles of individual samples [3], which should be less dependent on sequencing depth, biological replicates did generally not cluster together. This is in line with previous findings [3], which indicated that only largely skewed repertoire samples (like those from CLL

patients) allow separation according to immune status, while samples from immunized cohorts were not skewed enough and rather clustered according to library preparation protocol or sequencing machine than according to immunological status.

The B cell immune response to *Borrelia* is far from being completely understood. Even if our patients were diagnosed with acute infection, the level of VlsE-C6 reactive switched B cells were generally low. They showed even slightly lower or similar levels than cells induced by commonly encountered previous infections or vaccinations. Although the level of reactive B cells towards a single epitope (VlsE-C6) might not be directly comparable to the one towards whole pathogens or antigens, previous studies from the Lyme disease field have indicated that we might not expect drastic changes in peripheral B cell numbers upon acute infection [317,318]. Not only in mouse studies was it shown that *Borrelia* can actively influence the B cell immune response [289-292,303], but also in humans the B cell immune response seems to be rather of low level and different compared to other diseases [251].

5.2 Strengths and weaknesses of the approach used herein to identify *Borrelia* specific B cell repertoire signatures

Initial library preparation protocols were based on multiplex PCR using BIOMED2 primers [335] followed by adapter ligation. These were the standard Ion Torrent library preparation protocols at that time and had several disadvantages, which we tried to overcome by adapting the protocols to newer versions. The adapter ligation method required separation of the final library from unwanted side products by using E-gel. Because of differences in sizes of the individual molecules, retrieval of the complete library turned out to be rather problematic. Besides this, the standard reverse transcription protocol in combination with the high number of cycles (35) used in our multiplex PCR gave a disturbing side band, which had to be removed using Gel-extraction. This is enhancing the risk of sample cross contaminations, which we absolutely wanted to avoid when studying sequence overlap between samples. In order to make sure that each primer is working equally well in the PCR reaction, concentration of each primer would have needed to be optimized and internal controls should have been used for quantification of individual V-genes after amplification. This would have been a tremendous amount of work and since each repertoire might have a different sequence composition,

one cannot exclude that even after such extensive optimizations and especially at later PCR amplification steps the presence of certain V-genes might still influence amplification of others. Due to our own and other studies on barcoding techniques [364], we now know, that even when using only one primer pair, the PCR process is still a stochastic and heterogeneous process. Different strategies (including 5'-template switching [365], synthetic immune receptor repertoires for optimization of PCR conditions [366], resequencing of samples using different primers for amplification [367], emulsion PCR [368] and barcoding techniques [1]) have been developed to overcome the problems of unequal amplifications with multiplex PCR. Because of its simplicity and because it allows to get maximum quality reads, we decided to use a barcoding approach similar to the one that has been published by Vollmers et al. [1]. By labelling cDNA molecules with individual barcodes at the reverse transcription step and avoiding multiplex PCR, we are able to reduce PCR amplification bias and sequencing errors. To minimize library preparation or sequencing Chip effects, we made sure to always handle control and acute patient samples in parallel for each step of the experiment – meaning from RNA extraction through to the final sequencing run. Stringent precautions were also applied to avoid cross-sample contaminations. To have a more or less random distribution among samples from the different groups and timepoints, we also made sure to rotate the MID barcodes between samples each time libraries were prepared. By reducing the number of PCR cycles and by employing a reverse transcription protocol for high GC content mRNA molecules, we were able to get rid of the disturbing large unwanted side-band mentioned before. Although we were able to use a lower number of PCR cycles (19 in our case) as compared to the protocol of Vollmer et al. [1] (which used 27), the limiting step was actually not the amount of library needed for sequencing, but quantification on the bioanalyzer instrument. PCR approaches allowing quantification from lower amounts of library material might allow to further go down with PCR cycles and thereby give us an even less biased view into the B cell repertoire.

While manually generating phylogenetic trees of clones of interest including next generation sequencing data, we noticed that they still contained a large number of insertion and deletion errors. Those could be easily corrected by alignment to the single cell sequences. When assessing individual runs before and after cleanup, we were however able to see a reduction in errors, meaning that the used pipeline does its job. Insertion and deletion errors might persist in the final consensus sequences, because the Ion Torrent errors are not occurring randomly [369,370]. Since error correction with barcoding techniques is based on the most frequently occurring base at a certain position, systematic errors are very likely translated into the final consensus sequences [371] and could explain our

findings. Insertion and deletion errors can however be more easily identified and removed as compared to substitution errors and sequence assignments by IMGT [354] also takes these types of possible errors into account. With the present method, we are not able to control for errors occurring during the reverse transcription or second strand synthesis steps, nor are we be able to detect an error occurring in one of the very early cycles of the PCR used for amplification.

In order to get an idea about sequencing depth, we compared CDR3 amino acid overlap between sequencing, library and biological replicates. We concluded, that biological replicates are more reliable to get a maximum of information from the data. The overlap between biological replicates observed in the present study is comparable to the extent of overlap observed in other studies[1,12], confirming previous statements that different methods give rise to comparable results [12,372]. By comparing V and J gene distributions obtained with our next generation sequencing approach with that one obtained by randomly sorted single B cells, we were able to exclude large amplification biases. Similar to previously published data [12,373], IGHV3-23 and IGHV3-30 V-genes and IGHJ4 followed by IGHJ6 J-genes predominated within sequences obtained with both approaches. Although globally seen, there were no huge differences in the V and J gene usage in our datasets as compared to previously published ones slight discrepancies between the datasets were however observed. IGHV3-7 for example was more predominant in our datasets, while IGHJ6 was a bit less prominent as compared to those of other studies [12,373].

Since herein we analyzed a smaller number of sequences as compared to other studies [1,12], it makes sense that the degree of sample overlap observed here is closer to the range that they observe when only taking most abundant sequences into account. The ~20-30% CDR3 amino acid overlap between biological replicates indicates that our approach covers a significant part of the expanded repertoire. Also in line with what has previously been observed [12,14], we found that every donors' repertoire is rather unique. While ~10-30% CDR3 amino acid overlap was observed between samples from different timepoints, overlap between donors was not exceeding an average of 0.2%. Since the focus of the present study was the portion of the repertoire that is expanded in response to infection, the employed method should be perfectly fine. In the course of the study, we however found that overlapping clones might not be as expanded as initially hypothesized. Sequencing depth was however enough to see a difference in inter-donor IgG cluster overlap between our two groups. We were also able to find some single cell related clones back in the next generation sequencing data.

From our observations, we concluded that we are only sampling a snapshot of the bulk repertoire with deep sequencing but also only part of the antigen reacting B cell repertoire with the single cell approaches. Due to the heterogeneity and low level of *Borrelia* reactive B cells, future studies would need to extract sequences form much larger amounts of blood to be able to more closely assess clonal relatedness and evolution of *Borrelia* reactive B cells in the course of infection. As opposed to vaccination studies, in which flow cytometry [116], ELISPOT [67,135,136] and more recently also deep sequencing data [1,127,169,171,326,374] could confirm the plasmablast burst that generally occurs 6-7 days after the boost, the situation in acute or chronic infections might be more complicated. In addition to that, recent studies indicate that these largely expanded cells might not necessarily be the ones of interest for biomarker research, as lower level clones showed a higher degree of overlap between donors [171].

Our clustering analysis confirmed reports from other groups, who showed that it is possible to find a higher cluster overlap among individuals that encountered the same antigenic challenge [165,169,171,326]. These approaches allow to determine whether highly similar B cell clones are generated in different donors. We observed however, that very different B cell receptor sequences might interact with the same epitope and that the chance of finding similar B cell clones in the expanded repertoire of two different acute patients is rather low. Probably key residues and whole antibody structures rather than CDR3 sequences and common V and J genes are the determinants of antigen-binding. For this reason more sophisticated clustering approaches are required to extract sequences reactive with the same antigen. Bioinformatics tools that allow to extract common key mutated residues and structural features in a high throughput manner directly from deep sequencing data need to be developed. Due to the aforementioned problems with sample normalization, extraction of clones of an ongoing immune challenge will also be more reliable if general repertoire characteristics that could separate clones according to their time and mode of generation (like clonal expansions, mutation status, isotypes, phylogenetic tree structures, evolution of clones over the different timepoints, etc) are included into the analysis. The development of such a tool by far exceeds the capacities of this project. Publications of improved analysis pipelines for B cell repertoire data are exploding at the moment and probably some advanced tools will emerge soon.

Isolation of memory B cells of different specificities by *in vitro* single cell stimulation allowed us to confirm some of the problems we expected to have when using rather simple CDR3 clustering

approaches to identify related sequences in deep sequencing data. At low identity thresholds, the degree of clustering of single cell CDR3 amino acid sequences with deep sequencing data was rather dependent on sequence composition and similar values were obtained for the different donors. This indicates, that the degree of clustering of a certain sequence can depend on its composition and might not necessarily be due to the presence of a higher amount of closer clones in the repertoire. At lower levels of identity, sequences rich in tyrosine residues showed a higher tendency to be found back in the deep sequencing data. B cell repertoire signatures identified for Dengue [165] and influenza [169] viruses also showed such polytyrosine residues, raising the question of how specific antibodies with such properties are for a certain antigen. The physicochemical properties of tyrosine allows it to interact with a large variety of amino acids on the surface of antigens, which explains its abundance in antigen binding sites of antibodies [80,83]. Because of the lower abundance in the heavy chain CDR3s of antigen-experienced B cell subsets [73,81], sequences with polytyrosine residues were thought to be rather polyspecific, a finding that could not be confirmed by more recent studies [82]. Although polytyrosine residues were found in the CDR3s of several of our isolated single cell sequences, antibodies produced by the isolated cells generally gave positive results only for one of the several antigens tested, meaning that they are rather specific. Only one sequence which had two adjacent tyrosine residues (CDR3: ARADDSPSYYVNAFDL) showed reactivity to both *Borreli*a and measles virus. Very likely tyrosine residues are frequently used to mediate antigen contacts, but their precise position within a certain context is probably determining the specificity of these residues. Since the tested individuals were from the same region and were probably exposed to similar immune challenges in the past, it might not be surprising to find similar CDR3 sequence distributions and clustering tendencies within their memory repertoires. The only antigen for which we were able to compare truly negative with seropositive or acute donors was VlsE-C6. Also with this one we were not able to reliably define one clustering threshold that would allow to clearly separate acute from control groups. This is confirming, that each donor rather generates unique sequences in response to an immune challenge and that the rather individual nature of each binding mechanism complicates the determination of a global clustering scheme applicable to all antibody-antigen interactions. Isolation of antigen-reactive single B cell receptors and characterization of individual antibody-antigen interactions with x-ray crystallography remains the method of choice to better understand these phenomena.

Generally, it is also difficult to extract antigen-specific B cell signatures from the low numbers of antigen specific B cells isolated with single cell techniques. Our and previous studies [110-116] show

that reactive sequences can have very different V and J genes as well as CDR3 lengths. This makes them difficult to compare at the sequence level and more extensive analysis of the three dimensional antibody-antigen interactions would be required to completely understand which residues are important to determine antigen specificity. Mutation analysis of individual positions might be another strategy. The number of different clones per individual that are produced in response to a TT boost has for example been estimated to be in the range of 44-79 different clones per donor [116], a number that we are far from reaching with our single cell approach. When considering different donors, this number of possible interactions is expected to be even higher. We first need to determine the diversity and specificity of epitope reactive repertoires before being able to reliably answer the question of whether it is possible to use B cell repertoire signatures as diagnostic tools. In this regard, techniques combining single cell with deep sequencing approaches are crucial [123].

5.3 Comparison of methods allowing to isolate and characterize antigen-reactive B cells

One major goal of this study was to characterize B cell receptor sequences of clones that are naturally induced by *Borrelia* infection. For this reason, we explored different methods that allow to isolate antigen-reactive B cells directly from patient samples. One possibility we considered was immortalization of B cells [375]. This allows to reuse the same sample for different experiments. One could for example use one part of the expanded cells for sorting VlsE peptide positive B cells and at a later timepoint reuse the same sample for sorting with other *Borrelia* antigens. This would allow to dissect the B cell repertoire of a patient in great depth without the need of many patient samples. On the other hand, this approach has also already been used successfully to stimulate single cells and to isolate antigen-reactive B cells from patients [376]. In our hands, EBV immortalized B cell lines were however not very stable when kept for prolonged times in culture, which is why we abandoned this method. More recent immortalization techniques [377] seem very interesting as they allow to expand B cells expressing the B cell receptor on their surface but at the same time also secrete the soluble form of the receptor. This would allow to screen supernatants of single sorted cells for specific antibodies but one would also have the option to sort antigen-reactive B cells with tetramer staining approaches from such cultures.

Another attractive approach we started to explore was *in vitro* antigen-specific memory B cell stimulation. The idea was to enrich samples in B cell receptor sequences of interest prior to library preparation. From initial pilot experiments, we however concluded, that *in vitro* stimulation is a complex process that could be influenced by the activation status of the patient´s immune cells, the genetic background of the donor, peripheral blood immune cell subset composition and many other factors that are difficult to control for. Approaches, that pre-sort B cell subsets of interest and stimulate single isolated cells to produce antibodies followed by direct screening of supernatants by ELISA are beneficial over more general stimulation conditions [378,379]. These methods have the big advantage, that one can directly screen for reactivity of the B cell receptor towards the antigen, without the need for tedious cloning and expression of the antibodies. In addition to that, an increased amount of starting material due to the expansion of the single B cells should enhance the likelihood of a positive PCR result. Single cell approaches have also the advantage that one can get information about the whole B cell receptor, including corresponding heavy and light chain sequences. The stimulation protocol used in this study [355] has the advantage, that it allows to dissect the human memory B cell repertoire at the single cell level without any bias from prior pre-enrichment of cell subsets of interest. Although we negatively selected for IgG class switched B cells, we did not enrich for any other B cell marker allowing to get an idea about the level of antigen-specific B cells in the untouched IgG repertoire.

Combining *in vitro* single cell stimulation with screening against a whole group of antigens allows to determine antigen specificity of the single cells, thereby excluding the possibility of a positive ELISA result due to polyspecificity of the produced antibodies or other factors that might potentially influence ELISA results. Isolation of antigen-reactive memory B cells with tetramers [352] might be a less time consuming and cheaper approach, but setting the right gate is challenging especially if cell numbers are low and the amount of sample scarce. In addition to these problems, we also observed non-specific binding of the tetramers. It seems, that the *Borrelia* peptides bound to Neutravidin, but that not all the biotin binding sites of the tetramers were occupied. One possible explanation for this could be the presence of peptide aggregates. Some of the background noise could come from the binding of tetramers to $CD43^+$ B cells that are labeled with CD43-biotin from the MACS kit that we used to negatively enrich for B cells. Double staining of unmanipulated PBMC samples with peptide Streptavidin and Neutravidin tetramers in different colors will probably allow to reduce background noise to a maximum. Cloning and expression of the antibodies are crucial to separate truly peptide

reactive antibodies from non-specific ones. In theory, this step can be omitted when isolating antigen-reactive single cells by *in vitro* stimulation.

With the latter method, we were however only able to get a complete antibody sequence from a subset (40%) of positive wells. This was generally not due to the single cell PCR not amplifying all V genes, as 84% and 85% of the heavy and light chain PCRs were giving a PCR band visible on the gel. For 11% we did not get any PCR product at all. We however noticed, that in 18% of the wells we got a PCR product for both kappa and lambda light chains. In addition to that, for one of the antibodies, we got two different light chain sequences associated with one heavy chain sequence. For this reason, we conclude that also for this method it might still be important to clone the antibodies to confirm specificity. Since cell seeding follows Poisson distribution, it could potentially be that more than one cell has been seeded per well, which could be the reason why the Sanger sequencing did not work in some cases. To avoid confusion and if the antibody sequences are of interest, it might be important to consider high precision single cell sorting directly into the 384 well plate with a FACS machine instead of manual seeding. Although the rest of the sequences requires confirmation by cloning, we are quite confident, that at least for the heavy chain, the *Borrelia* peptide reactive sequence (CDR3: ARGTRDGQNPEFDY) isolated twice from Lyme 8 binds VlsE-C6.

Analysis of larger amounts of single cells are required to exclude that an insufficient sample coverage might be the reason for the low reproducibility between the two single cell isolation techniques finally used in this study. On the other hand, it is also possible that the two methods select for different B cell subsets.

Previous studies have shown, that antibodies might interact with antigens in their membrane bound form but do not necessarily need to also bind to the same antigen in their soluble form [380]. This might be one explanation for why IgM sequences isolated with labeled antigens of interest might seem rather non-specific when confirming binding with ELISA approaches [114]. Confirmation of antigen binding might therefore in that case be more directly assessed by employing competition assays [171]. I*n vitro* stimulation on the other hand might just activate specific B cell subsets thereby possibly not covering all *Borrelia* reactive B cells. Studies on hepatitis B vaccination also showed discrepancies between deep sequencing and ELISPOT data [126], which the authors explained by the possibility that the secreted antibodies might not have been specific enough to bind the antigen in the

ELISPOT assay. Deep sequencing on the other hand seemed sensitive enough to detect the B cell clones generated after the first vaccination boost.

The low frequency of clustering between sequences from our peptide enriched dataset originating from different donors is in line with other studies that assessed the degree of sequence overlap between datasets containing antigen enriched sequences [127,160]. We were able to confirm heterogeneity of VlsE-C6 reactive B cell receptor sequences with two different methods further supporting this notion.

5.4 Promises and challenges of studying the B cell repertoire of acute Lyme disease patients

Although public B cell repertoire signatures reactive to specific antigens have been discovered [127,165,169,171,326], their detection and validation generally still requires combination with approaches to isolate antigen-specific B cells. As illustrated by studies on anti-HIV broadly neutralizing antibodies [160,381] for example, it is rather difficult to extract antigen-specific signatures directly from deep sequencing data. Earlier single cell studies even suggested that it is rather not possible. Next generation sequencing now allows a deeper analysis of the B cell repertoire with increasing reports of sequences with overlapping features between donors. The diversity of the B cell repertoire is at the same time its beauty but also a nightmare in a sense, that it tremendously complicates data analysis. In this regard, Lyme disease especially in the European setting is an even worse nightmare. Different circulating *Borrelia* species [265] might express antigens of different structure. In addition to that, *Borrelia* are masters in changing protein expression [382]. Depending on the stage in which the patient is, different antigens might be expressed and predominate the B cell immune response. Not to forget the heterogeneous timepoints of infection.

Although upon acute infection, antibody secreting cell (ASC) numbers in peripheral blood also generally seem to peak around 7-8 days after onset of symptoms, responses can nevertheless vary greatly between donors [4]. Vaccination studies allow to analyze defined timepoints. Furthermore, individuals can be challenged with exactly the same antigen, leaving only the donors' background as variable. But even under these more "controlled" conditions, isolation of common clones seems rather challenging. With our approaches we were however able to detect a low level and heterogeneous

peripheral blood B cell immune response upon acute Lyme disease infection. The higher CDR3 amino acid overlap between acute patient samples from different timepoints is indicative of an ongoing B cell immune response at both the IgM and IgG levels. Furthermore, sequence clustering revealed a higher percentage overlap among acute Lyme disease patients' IgG repertoires as compared to those from healthy controls.

Pevious reports from the Lyme disease field indicate, that this disease might be associated with low and heterogeneous B cell immune responses and might explain the difficulties we had to isolate true *Borrelia* reactive B cell receptor sequences from our samples. In contrast to some older studies that showed an increase in peripheral B-lymphocytes in other acute bacterial infections [142,143], Lyme disease does not seem to be associated with such changes [317,318]. A recent RNAseq study of acute patients PBMCs found differences in the induction of B cell developmental pathways and calcium-induced T cell apoptosis by *Borrelia* as compared to other infectious agents [251]. These data generated further evidence for the involvement of a rather low proportion of B and T cells in peripheral blood in the acute Lyme disease condition as compared to other diseases. Analysis of immune factor levels within peripheral blood was able to separate acute patients into two groups, one with higher levels of T cell recruiting chemokine and inflammatory marker expression associated with lymphopenia and the development of antibodies against *Borrelia* and a second group in which B and T cell responses seem to be rather low [248], further supporting heterogeneous responses among patients and explaining possible different outcomes. Studies at the antibody level also described three different antibody responses towards *Borrelia* [304].

5.5 Possible effect of the tick bite on IgM expressing B cells

Both our flow cytometry and the deep sequencing data indicate, that a subset of individuals develop a strong B cell immune response to the tick bite, a phenomenon, that we did not observe to such a drastic degree in the acute Lyme disease patients. As in three out of the five analyzed donors with a recent tick bite we observed a strong and long lasting perturbation at the level of IgM expressing B cells, we think it might be interesting to analyze larger cohorts of donors with a recent tick bite to confirm this effect. We can envisage two possible explanations for the absence in acute Lyme disease. It could be, that *Borrelia* counteract such strong B cell immune responses similar to the effects observed in the mouse model [289-293,303], thereby possibly generating beneficial conditions for

the tick to take an undisturbed meal. Since this effect seems to occur already before the general transmission of *Borrelia* [184,185], another more plausible explanation would be, that *Borrelia* are rather transmitted in individuals with weaker responses towards ticks [383]. It is maybe interesting to mention in this context, that the third tick bite donor, which did not contain largely expanded clones in its IgM repertoire, showed a very high IgG sequence overlap with acute patients. This indicates that the response in this donor might be closer to a response occurring upon acute Lyme disease infection. As it is unnaturally quick and long-lasting, the real existence of this effect needs further validation. The lower diversity of the two affected repertoires might not necessarily be due to a B cell expansion but could also be due to lower B cell numbers in these individuals. This would support the possibility that naïve B cells disappeared from the blood and that other cells might have predominated these PBMC samples. We think that it is highly unlikely, that merely by chance the same sequence was overamplified in the different samples from the same donor. Also the mutation status of these sequences indicated, that they are quite distant from the germline, while this was generally not true for related sequences found in other donors. These findings indicate, that these mutations are probably true mutations and were not solely present due to difficulties of the Ion Torrent to sequence reads of that particular composition. In the expanded clone from Tick2, replacement mutations were mainly concentrated in the CDR2 region, which would make biologically sense.

5.6 Interesting individual cases

Two (Lyme2 and Lyme8) of the twelve acute Lyme disease patients analyzed in this study showed an increase in VlsE-C6 IgG antibody titer as determined by our in house peptide ELISA. One of these donors (Lyme8) has been analyzed with *in vitro* single B cell stimulation. This was the only case for which we were able to isolate the same B cell sequence from two different wells, making us confident that this sequence is really *Borrelia* specific. Since this donor was showing a drastic increase in antibody production over the sampling period, we would expect to detect expanded plasmablasts in these peripheral blood samples. In accordance with an ongoing response, this donor showed a quite high CDR3 amino acid overlap between samples from the different timepoints (IgG: 33%(T0)↔7%(T1), 6%(T1)↔8%(T2), 33%(T0)↔8%(T2) , IgM: 60%(T0)↔50%(T1), 53%(T1)↔52%(T2), 58%(T0)↔51%(T2)). A large number (123) of sequences from the deep sequencing data could be assigned to the clone isolated by *in vitro* single cell stimulation, which was dominated by sequences from the first sampling timepoint. This donor reported that the symptoms

started 4 days before the first blood draw, indicating that she is still at a very early stage of the disease. When integrating this information with the results obtained from phylogenetic tree analysis, we concluded that this clone might have been generated from memory B cells. All sequences of that clone were of IgG isotype and this tree had a quite long tree trunk, indicating that this clone might have started from an already mutated B cell. Interestingly we had difficulties to isolate *Borrelia* reactive memory B cells from this donor with our tetramer staining approach. We managed to isolate two peptide reactive sequences from the first timepoint sample. One of these could be found back in the next generation sequencing data. This clone was predominated by IgM sequences that could be found in all three timepoints analyzed, but interestingly one IgG sequence was found in the last timepoint. Since the trunk of this tree is also rather long, this clone has probably been generated from an IgM memory B cell that switched isotype at later timepoints of infection. Together these data indicate that this donors B cell immune response started from pre-existing memory B cells and might therefore have been rather plasma cell than memory B cell driven. These findings indicate, that we are isolating different B cell subsets with our two single cell approaches. Single cell clonal analysis are however in line with next generation sequencing data, indicating that a short term expansion of IgG B cells occurred at the first timepoint of sampling, combined with a rather long term perturbation at the IgM level. More general phylogenetic tree analysis of the deep sequencing data [77,384,385] from our patients will allow to classify clones into different categories. Repertoire data harbour a tremendous amount of valuable information that would allow to determine the stage and the nature of individual B cell immune responses. The possibility of differentiating a primary from a secondary immune response based on B cell repertoire data can be very helpful to support correct interpretation of serological results.

Although we cannot exclude, that Lyme4 was recently vaccinated against tetanus toxoid (TT), interestingly we detected a possible ongoing immune response against this antigen in this donor. This patient had problems to mount a robust class switched long term antibody response towards VlsE-C6, as the IgG antibody titres were declining with treatment over time. Besides erythema migrans, this patient also reported to have suffered from additional symptoms which did not completely disappear within the one month period sampled. Although CDC [386] is warning that TT might be transferred by insect bites, I am not aware of any study directly showing infection by *Clostridium tetani* via tick bites. The possibility is however not to be excluded. Although this patient might have been infected via another route, it could very likely be, that ticks are contaminated with spores that could be transmitted during blood feeding. This might be a topic worth investigating, as ticks are

outside in nature, where such spores could be present. This means, that it might be crucial to regularly check the TT immune status of risk cohorts - as should however theoretically anyhow be done – especially in the case of more elderly people [387]. Transmission via insect bites might also be one explanation for why tetanus infections are common in countries with high malaria incidences [388].

It might also be interesting to note, that Lyme6 had only symptoms for two weeks and stayed seronegative over the one month period of sampling. Quite some of the single cells isolated with tetramer staining from that donor could however be found back in the next generation sequencing data. Clones of both IgM and IgG isotypes could be retrieved in that way. The presence of memory B cells but absence of antibodies in this early stage patient is in line with mouse studies, that showed that the tick can have an effect on antibody production but not memory B cell formation [287,288]. This might be one explanation for the rather slow antibody response of patients towards *Borrelia*. Due to problems with non-specific binding, we can however not be completely sure that the isolated single cells are really VlsE-C6 reactive so that further confirmation of such a phenomenon is required. With *in vitro* single cell stimulation, we were however also able to identify VlsE-C6 reactive B cell clones in that donor, indicating that they might indeed be present in this repertoire.

A similar phenomenon was also observed in donor SP2. This was the donor from the seropositive group that showed the brightest staining with peptide tetramers. Initially this donor was assigned to the seronegative group. When reanalyzing the immunoblot after having observed this bright signal induced by tetramer staining, we however observed a faint IgG band reactive with the VlsE-mix. This donor was probably categorized as being seronegative, because reactivity was not clearly more positive as compared to the cut-off control of the immunoblot. When reassessing the donors' data sheets, we noticed that he reported to have had a tick bite with a red skin lesion 5 weeks before the blood draw. For these reasons, we shifted this donor later into the seropositive group. Interestingly, when preparing the B cells for the tetramer staining experiment, we observed an overnight proliferation of the B cells. The next morning way more cells were present in the tube as compared to the day before. This was however the only donor for which we observed such an effect. Interestingly 72% of the single sorted B cells from that donor were of IgA isotype, while this was the case for only 46% of the control cells. IgA anti-VlsE-C6 antibodies have been detected in early Lyme disease, but IgA only positive individuals have not been observed in that study[389]. Overall our results obtained with the peptide tetramer staining approach should be interpreted with caution. We

however repeatedly observed a little bit more staining in seropositive as compared to seronegative donors. In a mouse study from our laboratory, the same approach (using a different antigen and peptide) confirmed a brighter staining of B cells from immunized as compared to non-immunized mice. From this, we concluded that at least some of the stained B cells should be peptide reactive. We did however not observe any correlation between antibody levels and memory B cell staining. The same was also observed in other studies [390], meaning that this approach cannot be used to confirm specificity. Competition experiments [171] should have been performed to more reliably confirm specificity of the approach.

The last interesting case that we want to mention in this context is Lyme11. In this donor, we identified a CLL clone that interferred with our tetramer staining and that allowed us to understand some possible reasons for background noise in this type of assay. This donor showed a rather special response towards *Borrelia*, because the erythema migrans was present already for a whole year before Lyme disease was diagnosed and the patient entered our study. Although erythema migrans is thought to generally resolve even without antibiotics treatment, this seemed rather problematic in this patient. Upon antibiotics treatment, the erythema migrans however disappeared while other symptoms remained. This shows, that it might be important to consider other possible reasons – like a weakened immune response or other patient specific factors that might affect the immune response - than *Borrelia* themselves for the chronic course of Lyme disease. It was found, that patients with haematological malignancies for example more frequently required retreatment [218]. This indicates, that these patients might have more problems to combat the bacteria and would be in line with what we observed in this patient. It remains to be investigated whether the continuous symptoms in this case have been due to *Borrelia* or the underlying CLL condition.

6 CONCLUSION AND PERSPECTIVES

Own experiences with recruitment of patients confirmed, that problems with the diagnosis of acute Lyme disease are real. Although diagnosis might be easy for experts in the field, it might be rather problematic for general practitioners that only rarely encounter these patients. The rather personal nature of the B cell repertoire hampered the isolation of a clear *Borrelia* specific B cell repertoire signature. Our data however indicate that an ongoing B cell immune response is detectable in the peripheral blood of acute patients at both the IgM and IgG levels. The question to solve in future projects will be to determine in what exact way information contained in these data are useful to support diagnosis. With the present project, we are able to provide a first insight into the composition of B cell repertoires of acute Lyme disease patients. Since many different B cell receptors can interact with the same epitope, we came to the conclusion that we first need to determine the sizes, diversities and specificities of antigen-reactive repertoires before being able to reliably extract signatures from the deep sequencing data. Once candidate signatures are identified, quantitative real time PCR might be more sensitive to detect the clones of interest in patient samples than next generation sequencing. Based on our findings, we don't expect individual well defined clones to overlap between different patients. The signatures that we are looking for seem way more complex than initially imagined. We expect that more extensive repertoire analysis will lead to the extraction of a list of well-defined *Borrelia*-reactive B cell clones that might be separated from clones of other reactivities by the presence of key mutations at certain residues. Due to the rather dynamic and private nature of the B cell repertoire, the presence of certain combinations of clones will probably be the signatures that we are looking for (Figure 55). We expect, that mutation levels and isotype compositions of individual clones present within the repertoire as well as their phylogenetic tree relationships [77,384,385] will allow to classify patients into different categories. We expect that a closer analysis of these parameters will allow to distinguish ongoing B cell immune responses from past infections and primary from secondary immune responses. We would for example expect more isotype mixed clones that are closer to the germline in a primary immune response as compared to a secondary one. If we manage to match precisely defined clones with certain antigens, we might use the B cell repertoire to distinguish chronic infections from autoimmune responses. This could be very helpful to understand the underlying causes of "chronic Lyme disease". The B cell repertoire is a very attractive tool to be

explored in all cases for which direct detection of the pathogen within patient samples is not reliable. Due to its complex nature more extensive investigations are however required before we completely understand B cell immune responses at this level of detail. Bioinformatics and systems biology approaches are required to process the tremendous amount of information present in these data and will allow to extract only the relevant information. Before being able to develop sophisticated bioinformatics tools, one does however need to explore and understand the nature of these data. This is what we provide with the present project. Now will be the turn of the bioinformaticians to develop the necessary tools to more robustly validate our hypothesis (Figure 55).

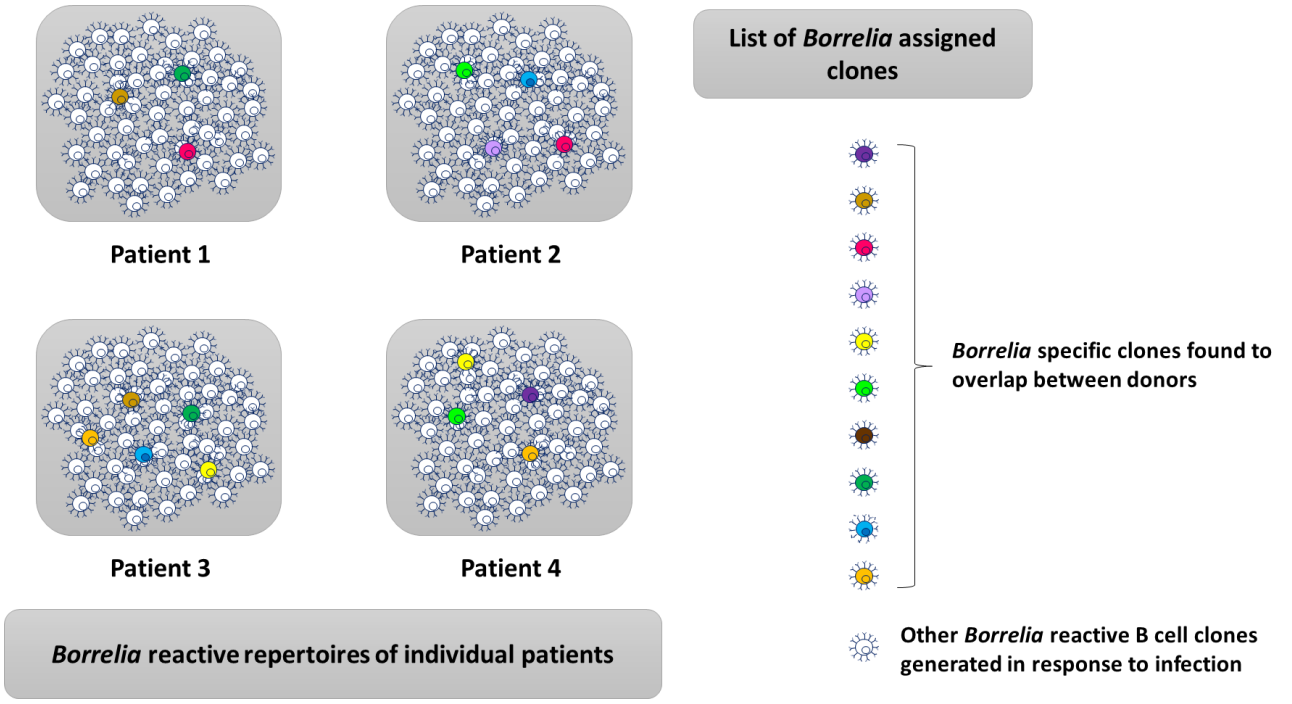


Figure 55: Schematic representation of how B cell repertoire data might be used in the future to support the diagnosis of acute Lyme disease. *From our data, we concluded that there will be rather a large number of different Borrelia reactive clones and that only a subset of overlapping clones will occur in each patient. Of those, again only a subset will be sampled when taking blood. More extensive analysis of Borrelia reactive B cell repertoires will be required to extract specific clones of interest. Future diagnostic tests based on B cell repertoire data will determine the presence of these identified clones within the repertoires of patients. The polyspecific nature of the B cell repertoire will probably require the presence of more than one of these clones for the tests to be reliable. Primary, secondary, ongoing and past immune responses can be distinguished by integrating more general repertoire characteristics of Borrelia reactive clones.*

7 REFERENCES

1. Vollmers C, Sit RV, Weinstein JA, Dekker CL, Quake SR (2013) Genetic measurement of memory B-cell recall using antibody repertoire sequencing. Proc Natl Acad Sci U S A 110: 13463-13468.

2. Bischof J, Ibrahim SM (2016) bcRep: R Package for Comprehensive Analysis of B Cell Receptor Repertoire Data. PLoS One 11: e0161569.

3. Greiff V, Bhat P, Cook SC, Menzel U, Kang W, et al. (2015) A bioinformatic framework for immune repertoire diversity profiling enables detection of immunological status. Genome Med 7: 49.

4. Carter MJ, Mitchell RM, Meyer Sauteur PM, Kelly DF, Truck J (2017) The Antibody-Secreting Cell Response to Infection: Kinetics and Clinical Applications. Front Immunol 8: 630.

5. Landry ML (2016) Immunoglobulin M for Acute Infection: True or False? Clin Vaccine Immunol 23: 540-545.

6. Sela-Culang I, Kunik V, Ofran Y (2013) The structural basis of antibody-antigen recognition. Front Immunol 4: 302.

7. Kabat EA (1982) Antibody diversity versus antibody complementarity. Pharmacol Rev 34: 23-38.

8. Nagasawa T (2006) Microenvironmental niches in the bone marrow required for B-cell development. Nat Rev Immunol 6: 107-116.

9. Nunez C, Nishimoto N, Gartland GL, Billips LG, Burrows PD, et al. (1996) B cells are generated throughout life in humans. J Immunol 156: 866-872.

10. Tonegawa S (1983) Somatic generation of antibody diversity. Nature 302: 575-581.

11. Xu JL, Davis MM (2000) Diversity in the CDR3 region of V(H) is sufficient for most antibody specificities. Immunity 13: 37-45.

12. Galson JD, Truck J, Fowler A, Munz M, Cerundolo V, et al. (2015) In-Depth Assessment of Within-Individual and Inter-Individual Variation in the B Cell Receptor Repertoire. Front Immunol 6: 531.

13. Jackson KJ, Kidd MJ, Wang Y, Collins AM (2013) The shape of the lymphocyte receptor repertoire: lessons from the B cell receptor. Front Immunol 4: 263.

14. Glanville J, Kuo TC, von Budingen HC, Guey L, Berka J, et al. (2011) Naive antibody gene-segment frequencies are heritable and unaltered by chronic lymphocyte ablation. Proc Natl Acad Sci U S A 108: 20066-20071.

15. DeKosky BJ, Lungu OI, Park D, Johnson EL, Charab W, et al. (2016) Large-scale sequence and structural comparisons of human naive and antigen-experienced antibody repertoires. Proc Natl Acad Sci U S A 113: E2636-2645.

16. Brezinschek HP, Foster SJ, Dorner T, Brezinschek RI, Lipsky PE (1998) Pairing of variable heavy and variable kappa chains in individual naive and memory B cells. J Immunol 160: 4762-4767.

17. Larimore K, McCormick MW, Robins HS, Greenberg PD (2012) Shaping of human germline IgH repertoires revealed by deep sequencing. J Immunol 189: 3221-3230.

18. Siegrist CA, Aspinall R (2009) B-cell responses to vaccination at the extremes of age. Nat Rev Immunol 9: 185-194.

19. Yanes RE, Gustafson CE, Weyand CM, Goronzy JJ (2017) Lymphocyte generation and population homeostasis throughout life. Semin Hematol 54: 33-38.

20. Gibson KL, Wu YC, Barnett Y, Duggan O, Vaughan R, et al. (2009) B-cell diversity decreases in old age and is correlated with poor health status. Aging Cell 8: 18-25.

21. Mroczek ES, Ippolito GC, Rogosch T, Hoi KH, Hwangpo TA, et al. (2014) Differences in the composition of the human antibody repertoire by B cell subsets in the blood. Front Immunol 5: 96.

22. Perez-Andres M, Paiva B, Nieto WG, Caraux A, Schmitz A, et al. (2010) Human peripheral blood B-cell compartments: a crossroad in B-cell traffic. Cytometry B Clin Cytom 78 Suppl 1: S47-60.

23. Giesecke C, Frolich D, Reiter K, Mei HE, Wirries I, et al. (2014) Tissue distribution and dependence of responsiveness of human antigen-specific memory B cells. J Immunol 192: 3091-3100.

24. Bagnara D, Squillario M, Kipling D, Mora T, Walczak AM, et al. (2015) A Reassessment of IgM Memory Subsets in Humans. J Immunol 195: 3716-3724.

25. Boyd SD, Marshall EL, Merker JD, Maniar JM, Zhang LN, et al. (2009) Measurement and clinical monitoring of human lymphocyte clonality by massively parallel VDJ pyrosequencing. Sci Transl Med 1: 12ra23.

26. Glanville J, Zhai W, Berka J, Telman D, Huerta G, et al. (2009) Precise determination of the diversity of a combinatorial antibody library gives insight into the human immunoglobulin repertoire. Proc Natl Acad Sci U S A 106: 20216-20221.

27. Arnaout R, Lee W, Cahill P, Honan T, Sparrow T, et al. (2011) High-resolution description of antibody heavy-chain repertoires in humans. PLoS One 6: e22365.

28. Seifert M, Kuppers R (2016) Human memory B cells. Leukemia 30: 2283-2292.

29. Nutt SL, Hodgkin PD, Tarlinton DM, Corcoran LM (2015) The generation of antibody-secreting plasma cells. Nat Rev Immunol 15: 160-171.

30. Vinuesa CG, Chang PP (2013) Innate B cell helpers reveal novel types of antibody responses. Nat Immunol 14: 119-126.

31. Mohanan D, Slutter B, Henriksen-Lacey M, Jiskoot W, Bouwstra JA, et al. (2010) Administration routes affect the quality of immune responses: A cross-sectional evaluation of particulate antigen-delivery systems. J Control Release 147: 342-349.

32. Tangye SG, Ma CS, Brink R, Deenick EK (2013) The good, the bad and the ugly - TFH cells in human health and disease. Nat Rev Immunol 13: 412-426.

33. Cerutti A, Puga I, Cols M (2012) New helping friends for B cells. Eur J Immunol 42: 1956-1968.

34. Egbuniwe IU, Karagiannis SN, Nestle FO, Lacy KE (2015) Revisiting the role of B cells in skin immune surveillance. Trends Immunol 36: 102-111.

35. Mauri C, Menon M (2017) Human regulatory B cells in health and disease: therapeutic potential. J Clin Invest 127: 772-779.

36. Mestas J, Hughes CC (2004) Of mice and not men: differences between mouse and human immunology. J Immunol 172: 2731-2738.

37. Bekeredjian-Ding I, Jego G (2009) Toll-like receptors--sentries in the B-cell response. Immunology 128: 311-323.

38. Griffin DO, Holodick NE, Rothstein TL (2011) Human B1 cells in umbilical cord and adult peripheral blood express the novel phenotype CD20+ CD27+ CD43+ CD70. J Exp Med 208: 67-80.

39. Perez-Andres M, Grosserichter-Wagener C, Teodosio C, van Dongen JJ, Orfao A, et al. (2011) The nature of circulating CD27+CD43+ B cells. J Exp Med 208: 2565-2566.

40. Griffin DO, Holodick NE, Rothstein TL (2011) Human B1 cells are CD3-: A reply to "A human equivalent of mouse B-1 cells?" and "The nature of circulating CD27+CD43+ B cells". J Exp Med 208: 2566-2569.

41. Covens K, Verbinnen B, Geukens N, Meyts I, Schuit F, et al. (2013) Characterization of proposed human B-1 cells reveals pre-plasmablast phenotype. Blood 121: 5176-5183.

42. Rothstein TL (2016) Natural Antibodies as Rheostats for Susceptibility to Chronic Diseases in the Aged. Front Immunol 7: 127.

43. Klein U, Rajewsky K, Kuppers R (1998) Human immunoglobulin (Ig)M+IgD+ peripheral blood B cells expressing the CD27 cell surface antigen carry somatically mutated variable region genes: CD27 as a general marker for somatically mutated (memory) B cells. J Exp Med 188: 1679-1689.

44. Anderson SM, Tomayko MM, Ahuja A, Haberman AM, Shlomchik MJ (2007) New markers for murine memory B cells that define mutated and unmutated subsets. J Exp Med 204: 2103-2114.

45. Weill JC, Weller S, Reynaud CA (2009) Human marginal zone B cells. Annu Rev Immunol 27: 267-285.

46. Kaminski DA, Wei C, Qian Y, Rosenberg AF, Sanz I (2012) Advances in human B cell phenotypic profiling. Front Immunol 3: 302.

47. Morbach H, Eichhorn EM, Liese JG, Girschick HJ (2010) Reference values for B cell subpopulations from infancy to adulthood. Clin Exp Immunol 162: 271-279.

48. Bonnefoy JY, Lecoanet-Henchoz S, Aubry JP, Gauchat JF, Graber P (1995) CD23 and B-cell activation. Curr Opin Immunol 7: 355-359.

49. Griffin DO, Rothstein TL (2012) Human b1 cell frequency: isolation and analysis of human b1 cells. Front Immunol 3: 122.

50. Kipps TJ, Stevenson FK, Wu CJ, Croce CM, Packham G, et al. (2017) Chronic lymphocytic leukaemia. Nat Rev Dis Primers 3: 17008.

51. Palanichamy A, Barnard J, Zheng B, Owen T, Quach T, et al. (2009) Novel human transitional B cell populations revealed by B cell depletion therapy. J Immunol 182: 5982-5993.

52. Bohnhorst JO, Bjorgan MB, Thoen JE, Natvig JB, Thompson KM (2001) Bm1-Bm5 classification of peripheral blood B cells reveals circulating germinal center founder cells in healthy individuals and disturbance in the B cell subpopulations in patients with primary Sjogren's syndrome. J Immunol 167: 3610-3618.

53. De Silva NS, Klein U (2015) Dynamics of B cells in germinal centres. Nat Rev Immunol 15: 137-148.

54. Victora GD, Nussenzweig MC (2012) Germinal centers. Annu Rev Immunol 30: 429-457.

55. Takemori T, Kaji T, Takahashi Y, Shimoda M, Rajewsky K (2014) Generation of memory B cells inside and outside germinal centers. Eur J Immunol 44: 1258-1264.

56. Smith KG, Light A, Nossal GJ, Tarlinton DM (1997) The extent of affinity maturation differs between the memory and antibody-forming cell compartments in the primary immune response. EMBO J 16: 2996-3006.

57. Nussenzweig V, Benacerraf B (1967) Antihapten antibody specificity and L chain type. J Exp Med 126: 727-743.

58. Eisen HN (2014) Affinity enhancement of antibodies: how low-affinity antibodies produced early in immune responses are followed by high-affinity antibodies later and in memory B-cell responses. Cancer Immunol Res 2: 381-392.

59. Yaari G, Vander Heiden JA, Uduman M, Gadala-Maria D, Gupta N, et al. (2013) Models of somatic hypermutation targeting and substitution based on synonymous mutations from high-throughput immunoglobulin sequencing data. Front Immunol 4: 358.

60. Seifert M, Kuppers R (2009) Molecular footprints of a germinal center derivation of human IgM+(IgD+)CD27+ B cells and the dynamics of memory B cell generation. J Exp Med 206: 2659-2669.

61. Jacob J, Kassir R, Kelsoe G (1991) In situ studies of the primary immune response to (4-hydroxy-3-nitrophenyl)acetyl. I. The architecture and dynamics of responding cell populations. J Exp Med 173: 1165-1175.

62. Jacob J, Kelsoe G (1992) In situ studies of the primary immune response to (4-hydroxy-3-nitrophenyl)acetyl. II. A common clonal origin for periarteriolar lymphoid sheath-associated foci and germinal centers. J Exp Med 176: 679-687.

63. Smith KG, Hewitson TD, Nossal GJ, Tarlinton DM (1996) The phenotype and fate of the antibody-forming cells of the splenic foci. Eur J Immunol 26: 444-448.

64. Paus D, Phan TG, Chan TD, Gardam S, Basten A, et al. (2006) Antigen recognition strength regulates the choice between extrafollicular plasma cell and germinal center B cell differentiation. J Exp Med 203: 1081-1091.

65. O'Connor BP, Vogel LA, Zhang W, Loo W, Shnider D, et al. (2006) Imprinting the fate of antigen-reactive B cells through the affinity of the B cell receptor. J Immunol 177: 7723-7732.

66. Phan TG, Paus D, Chan TD, Turner ML, Nutt SL, et al. (2006) High affinity germinal center B cells are actively selected into the plasma cell compartment. J Exp Med 203: 2419-2424.

67. Odendahl M, Mei H, Hoyer BF, Jacobi AM, Hansen A, et al. (2005) Generation of migratory antigen-specific plasma blasts and mobilization of resident plasma cells in a secondary immune response. Blood 105: 1614-1621.

68. Purtha WE, Tedder TF, Johnson S, Bhattacharya D, Diamond MS (2011) Memory B cells, but not long-lived plasma cells, possess antigen specificities for viral escape mutants. J Exp Med 208: 2599-2606.

69. Shinnakasu R, Inoue T, Kometani K, Moriyama S, Adachi Y, et al. (2016) Regulated selection of germinal-center cells into the memory B cell compartment. Nat Immunol 17: 861-869.

70. Baumgarth N (2013) How specific is too specific? B-cell responses to viral infections reveal the importance of breadth over depth. Immunol Rev 255: 82-94.

71. Taylor JJ, Pape KA, Jenkins MK (2012) A germinal center-independent pathway generates unswitched memory B cells early in the primary response. J Exp Med 209: 597-606.

72. Seifert M, Przekopowitz M, Taudien S, Lollies A, Ronge V, et al. (2015) Functional capacities of human IgM memory B cells in early inflammatory responses and secondary germinal center reactions. Proc Natl Acad Sci U S A 112: E546-555.

73. Wu YC, Kipling D, Leong HS, Martin V, Ademokun AA, et al. (2010) High-throughput immunoglobulin repertoire analysis distinguishes between human IgM memory and switched memory B-cell populations. Blood 116: 1070-1078.

74. Wu YC, Kipling D, Dunn-Walters DK (2011) The relationship between CD27 negative and positive B cell populations in human peripheral blood. Front Immunol 2: 81.

75. Berkowska MA, Driessen GJ, Bikos V, Grosserichter-Wagener C, Stamatopoulos K, et al. (2011) Human memory B cells originate from three distinct germinal center-dependent and -independent maturation pathways. Blood 118: 2150-2158.

76. Briney BS, Willis JR, McKinney BA, Crowe JE, Jr. (2012) High-throughput antibody sequencing reveals genetic evidence of global regulation of the naive and memory repertoires that extends across individuals. Genes Immun 13: 469-473.

77. Budeus B, Schweigle de Reynoso S, Przekopowitz M, Hoffmann D, Seifert M, et al. (2015) Complexity of the human memory B-cell compartment is determined by the versatility of clonal diversification in germinal centers. Proc Natl Acad Sci U S A 112: E5281-5289.

78. Oakes CC, Seifert M, Assenov Y, Gu L, Przekopowitz M, et al. (2016) DNA methylation dynamics during B cell maturation underlie a continuum of disease phenotypes in chronic lymphocytic leukemia. Nat Genet 48: 253-264.

79. Mian IS, Bradwell AR, Olson AJ (1991) Structure, function and properties of antibody binding sites. J Mol Biol 217: 133-151.

80. Fellouse FA, Wiesmann C, Sidhu SS (2004) Synthetic antibodies from a four-amino-acid code: a dominant role for tyrosine in antigen recognition. Proc Natl Acad Sci U S A 101: 12467-12472.

81. Clark LA, Ganesan S, Papp S, van Vlijmen HW (2006) Trends in antibody sequence changes during the somatic hypermutation process. J Immunol 177: 333-340.

82. Birtalan S, Zhang Y, Fellouse FA, Shao L, Schaefer G, et al. (2008) The intrinsic contributions of tyrosine, serine, glycine and arginine to the affinity and specificity of antibodies. J Mol Biol 377: 1518-1528.

83. Ramaraj T, Angel T, Dratz EA, Jesaitis AJ, Mumey B (2012) Antigen-antibody interface properties: composition, residue interactions, and features of 53 non-redundant structures. Biochim Biophys Acta 1824: 520-532.

84. Robin G, Sato Y, Desplancq D, Rochel N, Weiss E, et al. (2014) Restricted diversity of antigen binding residues of antibodies revealed by computational alanine scanning of 227 antibody-antigen complexes. J Mol Biol 426: 3729-3743.

85. Shlomchik MJ, Weisel F (2012) Germinal center selection and the development of memory B and plasma cells. Immunol Rev 247: 52-63.

86. Hou D, Ying T, Wang L, Chen C, Lu S, et al. (2016) Immune Repertoire Diversity Correlated with Mortality in Avian Influenza A (H7N9) Virus Infected Patients. Sci Rep 6: 33843.

87. Dal Porto JM, Haberman AM, Kelsoe G, Shlomchik MJ (2002) Very low affinity B cells form germinal centers, become memory B cells, and participate in secondary immune responses when higher affinity competition is reduced. J Exp Med 195: 1215-1221.

88. Shih TA, Meffre E, Roederer M, Nussenzweig MC (2002) Role of BCR affinity in T cell dependent antibody responses in vivo. Nat Immunol 3: 570-575.

89. Schwickert TA, Lindquist RL, Shakhar G, Livshits G, Skokos D, et al. (2007) In vivo imaging of germinal centres reveals a dynamic open structure. Nature 446: 83-87.

90. Schwickert TA, Alabyev B, Manser T, Nussenzweig MC (2009) Germinal center reutilization by newly activated B cells. J Exp Med 206: 2907-2914.

91. Kuraoka M, Schmidt AG, Nojima T, Feng F, Watanabe A, et al. (2016) Complex Antigens Drive Permissive Clonal Selection in Germinal Centers. Immunity 44: 542-552.

92. Childs LM, Baskerville EB, Cobey S (2015) Trade-offs in antibody repertoires to complex antigens. Philos Trans R Soc Lond B Biol Sci 370.

93. Berek C, Milstein C (1987) Mutation drift and repertoire shift in the maturation of the immune response. Immunol Rev 96: 23-41.

94. Berek C, Griffiths GM, Milstein C (1985) Molecular events during maturation of the immune response to oxazolone. Nature 316: 412-418.

95. Greiff V, Menzel U, Miho E, Weber C, Riedel R, et al. (2017) Systems Analysis Reveals High Genetic and Antigen-Driven Predetermination of Antibody Repertoires throughout B Cell Development. Cell Rep 19: 1467-1478.

96. Manivel V, Bayiroglu F, Siddiqui Z, Salunke DM, Rao KV (2002) The primary antibody repertoire represents a linked network of degenerate antigen specificities. J Immunol 169: 888-897.

97. Edwards BM, Barash SC, Main SH, Choi GH, Minter R, et al. (2003) The remarkable flexibility of the human antibody repertoire; isolation of over one thousand different antibodies to a single protein, BLyS. J Mol Biol 334: 103-118.

98. Baumgarth N (2000) A two-phase model of B-cell activation. Immunol Rev 176: 171-180.

99. Sperling R, Francus T, Siskind GW (1983) Degeneracy of antibody specificity. J Immunol 131: 882-885.

100. Richards FF, Konigsberg WH, Rosenstein RW, Varga JM (1975) On the specificity of antibodies. Science 187: 130-137.

101. Kranz LM, Diken M, Haas H, Kreiter S, Loquai C, et al. (2016) Systemic RNA delivery to dendritic cells exploits antiviral defence for cancer immunotherapy. Nature 534: 396-401.

102. Van Regenmortel MH (2012) Basic research in HIV vaccinology is hampered by reductionist thinking. Front Immunol 3: 194.

103. Van Regenmortel MH (2014) An Outdated Notion of Antibody Specificity is One of the Major Detrimental Assumptions of the Structure-Based Reverse Vaccinology Paradigm, Which Prevented It from Helping to Develop an Effective HIV-1 Vaccine. Front Immunol 5: 593.

104. Chandra A, Latov N, Wormser GP, Marques AR, Alaedini A (2011) Epitope mapping of antibodies to VlsE protein of Borrelia burgdorferi in post-Lyme disease syndrome. Clin Immunol 141: 103-110.

105. Wine Y, Horton AP, Ippolito GC, Georgiou G (2015) Serology in the 21st century: the molecular-level analysis of the serum antibody repertoire. Curr Opin Immunol 35: 89-97.

106. Lavinder JJ, Wine Y, Giesecke C, Ippolito GC, Horton AP, et al. (2014) Identification and characterization of the constituent human serum antibodies elicited by vaccination. Proc Natl Acad Sci U S A 111: 2259-2264.

107. Lee J, Boutz DR, Chromikova V, Joyce MG, Vollmers C, et al. (2016) Molecular-level analysis of the serum antibody repertoire in young adults before and after seasonal influenza vaccination. Nat Med 22: 1456-1464.

108. Zhang Y, Meyer-Hermann M, George LA, Figge MT, Khan M, et al. (2013) Germinal center B cells govern their own fate via antibody feedback. J Exp Med 210: 457-464.

109. Zhang Y, Garcia-Ibanez L, Toellner KM (2016) Regulation of germinal center B-cell differentiation. Immunol Rev 270: 8-19.

110. Coronella JA, Spier C, Welch M, Trevor KT, Stopeck AT, et al. (2002) Antigen-driven oligoclonal expansion of tumor-infiltrating B cells in infiltrating ductal carcinoma of the breast. J Immunol 169: 1829-1836.

111. Cheng J, Torkamani A, Grover RK, Jones TM, Ruiz DI, et al. (2011) Ectopic B-cell clusters that infiltrate transplanted human kidneys are clonal. Proc Natl Acad Sci U S A 108: 5560-5565.

112. Bradshaw EM, Orihuela A, McArdel SL, Salajegheh M, Amato AA, et al. (2007) A local antigen-driven humoral response is present in the inflammatory myopathies. J Immunol 178: 547-556.

113. Bardelli M, Alleri L, Angiolini F, Buricchi F, Tavarini S, et al. (2013) Ex vivo analysis of human memory B lymphocytes specific for A and B influenza hemagglutinin by polychromatic flow-cytometry. PLoS One 8: e70620.

114. Patil SU, Ogunniyi AO, Calatroni A, Tadigotla VR, Ruiter B, et al. (2015) Peanut oral immunotherapy transiently expands circulating Ara h 2-specific B cells with a homologous repertoire in unrelated subjects. J Allergy Clin Immunol 136: 125-134 e112.

115. Owens GP, Ritchie AM, Gilden DH, Burgoon MP, Becker D, et al. (2007) Measles virus-specific plasma cells are prominent in subacute sclerosing panencephalitis CSF. Neurology 68: 1815-1819.

116. Frolich D, Giesecke C, Mei HE, Reiter K, Daridon C, et al. (2010) Secondary immunization generates clonally related antigen-specific plasma cells and memory B cells. J Immunol 185: 3103-3110.

117. Doorenspleet ME, Klarenbeek PL, de Hair MJ, van Schaik BD, Esveldt RE, et al. (2014) Rheumatoid arthritis synovial tissue harbours dominant B-cell and plasma-cell clones associated with autoreactivity. Ann Rheum Dis 73: 756-762.

118. Wu X, Zhou T, Zhu J, Zhang B, Georgiev I, et al. (2011) Focused evolution of HIV-1 neutralizing antibodies revealed by structures and deep sequencing. Science 333: 1593-1602.

119. Lerner RA (2011) Rare antibodies from combinatorial libraries suggests an S.O.S. component of the human immunological repertoire. Mol Biosyst 7: 1004-1012.

120. Ma L, Wang X, Bi X, Yang J, Shi B, et al. (2017) Characteristics Peripheral Blood IgG and IgM Heavy Chain Complementarity Determining Region 3 Repertoire before and after Immunization with Recombinant HBV Vaccine. PLoS One 12: e0170479.

121. Goodwin S, McPherson JD, McCombie WR (2016) Coming of age: ten years of next-generation sequencing technologies. Nat Rev Genet 17: 333-351.

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

123. Busse CE, Czogiel I, Braun P, Arndt PF, Wardemann H (2014) Single-cell based high-throughput sequencing of full-length immunoglobulin heavy and light chain genes. Eur J Immunol 44: 597-603.

124. Murugan R, Imkeller K, Busse CE, Wardemann H (2015) Direct high-throughput amplification and sequencing of immunoglobulin genes from single human B cells. Eur J Immunol 45: 2698-2700.

125. Kesmir C, De Boer RJ (1999) A mathematical model on germinal center kinetics and termination. J Immunol 163: 2463-2469.

126. Galson JD, Truck J, Clutterbuck EA, Fowler A, Cerundolo V, et al. (2016) B-cell repertoire dynamics after sequential hepatitis B vaccination and evidence for cross-reactive B-cell activation. Genome Med 8: 68.

127. Galson JD, Clutterbuck EA, Truck J, Ramasamy MN, Munz M, et al. (2015) BCR repertoire sequencing: different patterns of B-cell activation after two Meningococcal vaccines. Immunol Cell Biol 93: 885-895.

128. Appanna R, Kg S, Xu MH, Toh YX, Velumani S, et al. (2016) Plasmablasts During Acute Dengue Infection Represent a Small Subset of a Broader Virus-specific Memory B Cell Pool. EBioMedicine 12: 178-188.

129. Fink K (2012) Origin and Function of Circulating Plasmablasts during Acute Viral Infections. Front Immunol 3: 78.

130. Slifka MK, Matloubian M, Ahmed R (1995) Bone marrow is a major site of long-term antibody production after acute viral infection. J Virol 69: 1895-1902.

131. Ahuja A, Anderson SM, Khalil A, Shlomchik MJ (2008) Maintenance of the plasma cell pool is independent of memory B cells. Proc Natl Acad Sci U S A 105: 4802-4807.

132. Halliley JL, Tipton CM, Liesveld J, Rosenberg AF, Darce J, et al. (2015) Long-Lived Plasma Cells Are Contained within the CD19(-)CD38(hi)CD138(+) Subset in Human Bone Marrow. Immunity 43: 132-145.

133. Bachmann MF, Kundig TM, Odermatt B, Hengartner H, Zinkernagel RM (1994) Free recirculation of memory B cells versus antigen-dependent differentiation to antibody-forming cells. J Immunol 153: 3386-3397.

134. Amanna IJ, Slifka MK (2010) Mechanisms that determine plasma cell lifespan and the duration of humoral immunity. Immunol Rev 236: 125-138.

135. Blanchard-Rohner G, Pulickal AS, Jol-van der Zijde CM, Snape MD, Pollard AJ (2009) Appearance of peripheral blood plasma cells and memory B cells in a primary and secondary immune response in humans. Blood 114: 4998-5002.

136. Clutterbuck EA, Salt P, Oh S, Marchant A, Beverley P, et al. (2006) The kinetics and phenotype of the human B-cell response following immunization with a heptavalent pneumococcal-CRM conjugate vaccine. Immunology 119: 328-337.

137. Ten Boekel E, Siegert CE, Vrielink GJ, Van Dam VC, Ceelen A, et al. (2007) Analyses of CD27++ plasma cells in peripheral blood from patients with bacterial infections and patients with serum antinuclear antibodies. J Clin Immunol 27: 467-476.

138. Harada Y, Kawano MM, Huang N, Mahmoud MS, Lisukov IA, et al. (1996) Identification of early plasma cells in peripheral blood and their clinical significance. Br J Haematol 92: 184-191.

139. Lane HC, Volkman DJ, Whalen G, Fauci AS (1981) In vitro antigen-induced, antigen-specific antibody production in man. Specific and polyclonal components, kinetics, and cellular requirements. J Exp Med 154: 1043-1057.

140. Czerkinsky C, Nilsson LA, Ouchterlony O, Tarkowski A, Gretzer C (1984) Detection of single antibody-secreting cells generated after in vitro antigen-induced stimulation of human peripheral blood lymphocytes. Scand J Immunol 19: 575-579.

141. Uytdehaag FG, Osterhaus AD, Loggen HG, Bakker RH, van Asten JA, et al. (1983) Induction of antigen-specific antibody response in human peripheral blood lymphocytes in vitro by a dog kidney cell vaccine against rabies virus (DKCV). J Immunol 131: 1234-1239.

142. Niklasson PM, Williams RC, Jr. (1974) Studies of peripheral blood T-and B-lymphocytes in acute infections. Infect Immun 9: 1-7.

143. Thorley JD, Smith JW, Luby JP, Sanford JP (1977) Peripheral blood lymphocyte response to acute infections in humans. Infect Immun 16: 110-114.

144. Crotty S, Felgner P, Davies H, Glidewell J, Villarreal L, et al. (2003) Cutting edge: long-term B cell memory in humans after smallpox vaccination. J Immunol 171: 4969-4973.

145. Pinna D, Corti D, Jarrossay D, Sallusto F, Lanzavecchia A (2009) Clonal dissection of the human memory B-cell repertoire following infection and vaccination. Eur J Immunol 39: 1260-1270.

146. Mei HE, Yoshida T, Sime W, Hiepe F, Thiele K, et al. (2009) Blood-borne human plasma cells in steady state are derived from mucosal immune responses. Blood 113: 2461-2469.

147. Goidl EA, Paul WE, Siskind GW, Benacerraf B (1968) The effect of antigen dose and time after immunization on the amount and affinity of anti-hapten antibody. J Immunol 100: 371-375.

148. McIntyre D, Zuckerman NS, Field M, Mehr R, Stott DI (2014) The V(H) repertoire and clonal diversification of B cells in inflammatory myopathies. Eur J Immunol 44: 585-596.

149. Dorner T, Lipsky PE (2004) Correlation of circulating CD27high plasma cells and disease activity in systemic lupus erythematosus. Lupus 13: 283-289.

150. Jacobi AM, Odendahl M, Reiter K, Bruns A, Burmester GR, et al. (2003) Correlation between circulating CD27high plasma cells and disease activity in patients with systemic lupus erythematosus. Arthritis Rheum 48: 1332-1342.

151. Liao HX, Lynch R, Zhou T, Gao F, Alam SM, et al. (2013) Co-evolution of a broadly neutralizing HIV-1 antibody and founder virus. Nature 496: 469-476.

152. von Budingen HC, Kuo TC, Sirota M, van Belle CJ, Apeltsin L, et al. (2012) B cell exchange across the blood-brain barrier in multiple sclerosis. J Clin Invest 122: 4533-4543.

153. Ferraro AJ, Drayson MT, Savage CO, MacLennan IC (2008) Levels of autoantibodies, unlike antibodies to all extrinsic antigen groups, fall following B cell depletion with Rituximab. Eur J Immunol 38: 292-298.

154. Slocombe T, Brown S, Miles K, Gray M, Barr TA, et al. (2013) Plasma cell homeostasis: the effects of chronic antigen stimulation and inflammation. J Immunol 191: 3128-3138.

155. Han S, Zheng B, Dal Porto J, Kelsoe G (1995) In situ studies of the primary immune response to (4-hydroxy-3-nitrophenyl)acetyl. IV. Affinity-dependent, antigen-driven B cell apoptosis in germinal centers as a mechanism for maintaining self-tolerance. J Exp Med 182: 1635-1644.

156. Shokat KM, Goodnow CC (1995) Antigen-induced B-cell death and elimination during germinal-centre immune responses. Nature 375: 334-338.

157. Pulendran B, Kannourakis G, Nouri S, Smith KG, Nossal GJ (1995) Soluble antigen can cause enhanced apoptosis of germinal-centre B cells. Nature 375: 331-334.

158. Allen D, Simon T, Sablitzky F, Rajewsky K, Cumano A (1988) Antibody engineering for the analysis of affinity maturation of an anti-hapten response. EMBO J 7: 1995-2001.

159. Boersch-Supan ME, Agarwal S, White-Scharf ME, Imanishi-Kari T (1985) Heavy chain variable region. Multiple gene segments encode anti-4-(hydroxy-3-nitro-phenyl)acetyl idiotypic antibodies. J Exp Med 161: 1272-1292.

160. Zhou T, Zhu J, Wu X, Moquin S, Zhang B, et al. (2013) Multidonor analysis reveals structural elements, genetic determinants, and maturation pathway for HIV-1 neutralization by VRC01-class antibodies. Immunity 39: 245-258.

161. Pappas L, Foglierini M, Piccoli L, Kallewaard NL, Turrini F, et al. (2014) Rapid development of broadly influenza neutralizing antibodies through redundant mutations. Nature 516: 418-422.

162. Liu M, Yang G, Wiehe K, Nicely NI, Vandergrift NA, et al. (2015) Polyreactivity and autoreactivity among HIV-1 antibodies. J Virol 89: 784-798.

163. Yamagami J, Payne AS, Kacir S, Ishii K, Siegel DL, et al. (2010) Homologous regions of autoantibody heavy chain complementarity-determining region 3 (H-CDR3) in patients with pemphigus cause pathogenicity. J Clin Invest 120: 4111-4117.

164. Cameron EM, Spencer S, Lazarini J, Harp CT, Ward ES, et al. (2009) Potential of a unique antibody gene signature to predict conversion to clinically definite multiple sclerosis. J Neuroimmunol 213: 123-130.

165. Parameswaran P, Liu Y, Roskin KM, Jackson KK, Dixit VP, et al. (2013) Convergent antibody signatures in human dengue. Cell Host Microbe 13: 691-700.

166. Balakrishnan T, Bela-Ong DB, Toh YX, Flamand M, Devi S, et al. (2011) Dengue virus activates polyreactive, natural IgG B cells after primary and secondary infection. PLoS One 6: e29430.

167. Hou D, Chen C, Seely EJ, Chen S, Song Y (2016) High-Throughput Sequencing-Based Immune Repertoire Study during Infectious Disease. Front Immunol 7: 336.

168. Galson JD, Pollard AJ, Truck J, Kelly DF (2014) Studying the antibody repertoire after vaccination: practical applications. Trends Immunol 35: 319-331.

169. Jackson KJ, Liu Y, Roskin KM, Glanville J, Hoh RA, et al. (2014) Human responses to influenza vaccination show seroconversion signatures and convergent antibody rearrangements. Cell Host Microbe 16: 105-114.

170. Truck J, Ramasamy MN, Galson JD, Rance R, Parkhill J, et al. (2015) Identification of antigen-specific B cell receptor sequences using public repertoire analysis. J Immunol 194: 252-261.

171. Galson JD, Truck J, Fowler A, Clutterbuck EA, Munz M, et al. (2015) Analysis of B Cell Repertoire Dynamics Following Hepatitis B Vaccination in Humans, and Enrichment of Vaccine-specific Antibody Sequences. EBioMedicine 2: 2070-2079.

172. Hubalek Z (2009) Epidemiology of lyme borreliosis. Curr Probl Dermatol 37: 31-50.

173. Marques AR (2015) Laboratory diagnosis of Lyme disease: advances and challenges. Infect Dis Clin North Am 29: 295-307.

174. Stanek G, Wormser GP, Gray J, Strle F (2012) Lyme borreliosis. Lancet 379: 461-473.

175. Steen CJ, Carbonaro PA, Schwartz RA (2004) Arthropods in dermatology. J Am Acad Dermatol 50: 819-842, quiz 842-814.

176. Hopla CE, Durden LA, Keirans JE (1994) Ectoparasites and classification. Rev Sci Tech 13: 985-1017.

177. Wikel SK (1999) Tick modulation of host immunity: an important factor in pathogen transmission. Int J Parasitol 29: 851-859.

178. Brossard M, Wikel SK (2004) Tick immunobiology. Parasitology 129 Suppl: S161-176.

179. Hovius JW, Levi M, Fikrig E (2008) Salivating for knowledge: potential pharmacological agents in tick saliva. PLoS Med 5: e43.

180. Kotal J, Langhansova H, Lieskovska J, Andersen JF, Francischetti IM, et al. (2015) Modulation of host immunity by tick saliva. J Proteomics 128: 58-68.

181. Ribeiro JM, Mather TN, Piesman J, Spielman A (1987) Dissemination and salivary delivery of Lyme disease spirochetes in vector ticks (Acari: Ixodidae). J Med Entomol 24: 201-205.

182. Dunham-Ems SM, Caimano MJ, Pal U, Wolgemuth CW, Eggers CH, et al. (2009) Live imaging reveals a biphasic mode of dissemination of Borrelia burgdorferi within ticks. J Clin Invest 119: 3652-3665.

183. Sultan SZ, Manne A, Stewart PE, Bestor A, Rosa PA, et al. (2013) Motility is crucial for the infectious life cycle of Borrelia burgdorferi. Infect Immun 81: 2012-2021.

184. Crippa M, Rais O, Gern L (2002) Investigations on the mode and dynamics of transmission and infectivity of Borrelia burgdorferi sensu stricto and Borrelia afzelii in Ixodes ricinus ticks. Vector Borne Zoonotic Dis 2: 3-9.

185. Kahl O, Janetzki-Mittmann C, Gray JS, Jonas R, Stein J, et al. (1998) Risk of infection with Borrelia burgdorferi sensu lato for a host in relation to the duration of nymphal Ixodes ricinus feeding and the method of tick removal. Zentralbl Bakteriol 287: 41-52.

186. Nigrovic LE, Thompson KM (2007) The Lyme vaccine: a cautionary tale. Epidemiol Infect 135: 1-8.

187. Piesman J, Hojgaard A (2012) Protective value of prophylactic antibiotic treatment of tick bite for Lyme disease prevention: an animal model. Ticks Tick Borne Dis 3: 193-196.

188. Piesman J, Hojgaard A, Ullmann AJ, Dolan MC (2014) Efficacy of an experimental azithromycin cream for prophylaxis of tick-transmitted lyme disease spirochete infection in a murine model. Antimicrob Agents Chemother 58: 348-351.

189. Warshafsky S, Lee DH, Francois LK, Nowakowski J, Nadelman RB, et al. (2010) Efficacy of antibiotic prophylaxis for the prevention of Lyme disease: an updated systematic review and meta-analysis. J Antimicrob Chemother 65: 1137-1144.

190. Lavik J-P (2012) Intravital microscopy of Borrelia burgdorferi: delineation of dissemination kinetics and persistence within murine skin. Thesis and Dissertations 357 http://utdrutoledoedu/theses-dissertations/357.

191. Petzke M, Schwartz I (2015) Borrelia burgdorferi Pathogenesis and the Immune Response. Clin Lab Med 35: 745-764.

192. Wilhelmsson P, Fryland L, Lindblom P, Sjowall J, Ahlm C, et al. (2016) A prospective study on the incidence of Borrelia burgdorferi sensu lato infection after a tick bite in Sweden and on the Aland Islands, Finland (2008-2009). Ticks Tick Borne Dis 7: 71-79.

193. Koedel U, Fingerle V, Pfister HW (2015) Lyme neuroborreliosis-epidemiology, diagnosis and management. Nat Rev Neurol 11: 446-456.

194. Hansen K, Lebech AM (1992) The clinical and epidemiological profile of Lyme neuroborreliosis in Denmark 1985-1990. A prospective study of 187 patients with Borrelia burgdorferi specific intrathecal antibody production. Brain 115 (Pt 2): 399-423.

195. Reik L, Jr., Burgdorfer W, Donaldson JO (1986) Neurologic abnormalities in Lyme disease without erythema chronicum migrans. Am J Med 81: 73-78.

196. Hofhuis A, Herremans T, Notermans DW, Sprong H, Fonville M, et al. (2013) A prospective study among patients presenting at the general practitioner with a tick bite or erythema migrans in The Netherlands. PLoS One 8: e64361.

197. Feder HM, Jr., Whitaker DL (1995) Misdiagnosis of erythema migrans. Am J Med 99: 412-419.

198. Stanek G, Strle F (2003) Lyme borreliosis. Lancet 362: 1639-1647.

199. Wormser GP, Dattwyler RJ, Shapiro ED, Halperin JJ, Steere AC, et al. (2006) The clinical assessment, treatment, and prevention of lyme disease, human granulocytic anaplasmosis, and babesiosis: clinical practice guidelines by the Infectious Diseases Society of America. Clin Infect Dis 43: 1089-1134.

200. Godar DA, Laniosz V, Wetter DA (2015) Lyme disease update for the general dermatologist. Am J Clin Dermatol 16: 5-18.

201. Stanek G, Fingerle V, Hunfeld KP, Jaulhac B, Kaiser R, et al. (2011) Lyme borreliosis: clinical case definitions for diagnosis and management in Europe. Clin Microbiol Infect 17: 69-79.

202. Cutler SJ, Rudenko N, Golovchenko M, Cramaro WJ, Kirpach J, et al. (2017) Diagnosing Borreliosis. Vector Borne Zoonotic Dis 17: 2-11.

203. Wilske B (2003) Diagnosis of lyme borreliosis in europe. Vector Borne Zoonotic Dis 3: 215-227.

204. DeBiasi RL (2014) A concise critical analysis of serologic testing for the diagnosis of lyme disease. Curr Infect Dis Rep 16: 450.

205. Moore A, Nelson C, Molins C, Mead P, Schriefer M (2016) Current Guidelines, Common Clinical Pitfalls, and Future Directions for Laboratory Diagnosis of Lyme Disease, United States. Emerg Infect Dis 22.

206. Miraglia CM (2016) A Review of the Centers for Disease Control and Prevention's Guidelines for the Clinical Laboratory Diagnosis of Lyme Disease. J Chiropr Med 15: 272-280.

207. Theel ES (2016) The Past, Present, and (Possible) Future of Serologic Testing for Lyme Disease. J Clin Microbiol 54: 1191-1196.

208. Lantos PM, Branda JA, Boggan JC, Chudgar SM, Wilson EA, et al. (2015) Poor Positive Predictive Value of Lyme Disease Serologic Testing in an Area of Low Disease Incidence. Clin Infect Dis 61: 1374-1380.

209. Fallon BA, Pavlicova M, Coffino SW, Brenner C (2014) A comparison of lyme disease serologic test results from 4 laboratories in patients with persistent symptoms after antibiotic treatment. Clin Infect Dis 59: 1705-1710.

210. Dattwyler RJ, Arnaboldi PM (2014) Comparison of lyme disease serologic assays and lyme specialty laboratories. Clin Infect Dis 59: 1711-1713.

211. Leeflang MM, Ang CW, Berkhout J, Bijlmer HA, Van Bortel W, et al. (2016) The diagnostic accuracy of serological tests for Lyme borreliosis in Europe: a systematic review and meta-analysis. BMC Infect Dis 16: 140.

212. van Gorkom T, Kremer K, Voet W, Notermans DW, Vlaminckx BJM, et al. (2017) Disagreement between the results from three commercial tests for the detection of Borrelia-specific serum antibodies in the Netherlands associated with antibiotic treatment for Lyme borreliosis: a retrospective study. Eur J Clin Microbiol Infect Dis.

213. Lantos PM (2015) Chronic Lyme disease. Infect Dis Clin North Am 29: 325-340.

214. Oliveira CR, Shapiro ED (2015) Update on persistent symptoms associated with Lyme disease. Curr Opin Pediatr 27: 100-104.

215. Maloney EL (2016) Controversies in Persistent (Chronic) Lyme Disease. J Infus Nurs 39: 369-375.

216. Fallon BA, Petkova E, Keilp JG, Britton CB (2012) A reappraisal of the u.s. Clinical trials of post-treatment lyme disease syndrome. Open Neurol J 6: 79-87.

217. Stricker RB (2007) Counterpoint: long-term antibiotic therapy improves persistent symptoms associated with lyme disease. Clin Infect Dis 45: 149-157.

218. Maraspin V, Ruzic-Sabljic E, Lusa L, Strle F (2015) Course and outcome of Early Lyme borreliosis in patients with hematological malignancies. Clin Infect Dis 61: 427-431.

219. Nadelman RB, Wormser GP (2007) Reinfection in patients with Lyme disease. Clin Infect Dis 45: 1032-1038.

220. Nadelman RB, Hanincova K, Mukherjee P, Liveris D, Nowakowski J, et al. (2012) Differentiation of reinfection from relapse in recurrent Lyme disease. N Engl J Med 367: 1883-1890.

221. Khatchikian CE, Nadelman RB, Nowakowski J, Schwartz I, Wormser GP, et al. (2014) Evidence for strain-specific immunity in patients treated for early lyme disease. Infect Immun 82: 1408-1413.

222. Aguero-Rosenfeld ME, Wang G, Schwartz I, Wormser GP (2005) Diagnosis of lyme borreliosis. Clin Microbiol Rev 18: 484-509.

223. Strle K, Stupica D, Drouin EE, Steere AC, Strle F (2014) Elevated levels of IL-23 in a subset of patients with post-lyme disease symptoms following erythema migrans. Clin Infect Dis 58: 372-380.

224. Marques A, Telford SR, 3rd, Turk SP, Chung E, Williams C, et al. (2014) Xenodiagnosis to detect Borrelia burgdorferi infection: a first-in-human study. Clin Infect Dis 58: 937-945.

225. Telford SR, 3rd, Hu LT, Marques A (2014) Is there a place for xenodiagnosis in the clinic? Expert Rev Anti Infect Ther 12: 1307-1310.

226. Bockenstedt LK, Radolf JD (2014) Xenodiagnosis for posttreatment Lyme disease syndrome: resolving the conundrum or adding to it? Clin Infect Dis 58: 946-948.

227. Lantos PM, Auwaerter PG, Wormser GP (2014) A systematic review of Borrelia burgdorferi morphologic variants does not support a role in chronic Lyme disease. Clin Infect Dis 58: 663-671.

228. Bockenstedt LK, Gonzalez DG, Haberman AM, Belperron AA (2012) Spirochete antigens persist near cartilage after murine Lyme borreliosis therapy. J Clin Invest 122: 2652-2660.

229. Wormser GP, Nadelman RB, Schwartz I (2012) The amber theory of Lyme arthritis: initial description and clinical implications. Clin Rheumatol 31: 989-994.

230. Kuenzle S, von Budingen HC, Meier M, Harrer MD, Urich E, et al. (2007) Pathogen specificity and autoimmunity are distinct features of antigen-driven immune responses in neuroborreliosis. Infect Immun 75: 3842-3847.

231. Suchanek G, Kristoferitsch W, Stanek G, Bernheimer H (1986) Anti-myelin antibodies in cerebrospinal fluid and serum of patients with meningopolyneuritis Garin-Bujadoux-Bannwarth and other neurological diseases. Zentralbl Bakteriol Mikrobiol Hyg A 263: 160-168.

232. Baig S, Olsson T, Hojeberg B, Link H (1991) Cells secreting antibodies to myelin basic protein in cerebrospinal fluid of patients with Lyme neuroborreliosis. Neurology 41: 581-587.

233. Kaiser R (1995) Intrathecal immune response in patients with neuroborreliosis: specificity of antibodies for neuronal proteins. J Neurol 242: 319-325.

234. Girouard L, Laux DC, Jindal S, Nelson DR (1993) Immune recognition of human Hsp60 by Lyme disease patient sera. Microb Pathog 14: 287-297.

235. Sigal LH, Williams S (1997) A monoclonal antibody to Borrelia burgdorferi flagellin modifies neuroblastoma cell neuritogenesis in vitro: a possible role for autoimmunity in the neuropathy of Lyme disease. Infect Immun 65: 1722-1728.

236. Eiffert H, Karsten A, Ritter K, Ohlenbusch A, Schlott T, et al. (2005) Autoantibodies to human manganese superoxide dismutase (MnSOD) in children with facial palsy due to neuroborreliosis. Neuropediatrics 36: 386-388.

237. Goebel KM, Krause A, Neurath F (1988) Acquired transient autoimmune reactions in Lyme arthritis: correlation between rheumatoid factor and disease activity. Scand J Rheumatol Suppl 75: 314-317.

238. Mackworth-Young CG, Harris EN, Steere AC, Rizvi F, Malawista SE, et al. (1988) Anticardiolipin antibodies in Lyme disease. Arthritis Rheum 31: 1052-1056.

239. Garcia Monco JC, Wheeler CM, Benach JL, Furie RA, Lukehart SA, et al. (1993) Reactivity of neuroborreliosis patients (Lyme disease) to cardiolipin and gangliosides. J Neurol Sci 117: 206-214.

240. Ghosh S, Seward R, Costello CE, Stollar BD, Huber BT (2006) Autoantibodies from synovial lesions in chronic, antibiotic treatment-resistant Lyme arthritis bind cytokeratin-10. J Immunol 177: 2486-2494.

241. Crowley JT, Strle K, Drouin EE, Pianta A, Arvikar SL, et al. (2016) Matrix metalloproteinase-10 is a target of T and B cell responses that correlate with synovial pathology in patients with antibiotic-refractory Lyme arthritis. J Autoimmun 69: 24-37.

242. Schluesener HJ, Martin R, Sticht-Groh V (1989) Autoimmunity in Lyme disease: molecular cloning of antigens recognized by antibodies in the cerebrospinal fluid. Autoimmunity 2: 323-330.

243. Kratz A, Harding MW, Craft J, Mackworth-Young CG, Handschumacher RE (1992) Autoantibodies against cyclophilin in systemic lupus erythematosus and Lyme disease. Clin Exp Immunol 90: 422-427.

244. Crowley JT, Drouin EE, Pianta A, Strle K, Wang Q, et al. (2015) A Highly Expressed Human Protein, Apolipoprotein B-100, Serves as an Autoantigen in a Subgroup of Patients With Lyme Disease. J Infect Dis 212: 1841-1850.

245. Pianta A, Drouin EE, Crowley JT, Arvikar S, Strle K, et al. (2015) Annexin A2 is a target of autoimmune T and B cell responses associated with synovial fibroblast proliferation in patients with antibiotic-refractory Lyme arthritis. Clin Immunol 160: 336-341.

246. Greco TP, Jr., Conti-Kelly AM, Greco TP (2011) Antiphospholipid antibodies in patients with purported 'chronic Lyme disease'. Lupus 20: 1372-1377.

247. Londono D, Cadavid D, Drouin EE, Strle K, McHugh G, et al. (2014) Antibodies to endothelial cell growth factor and obliterative microvascular lesions in the synovium of patients with antibiotic-refractory lyme arthritis. Arthritis Rheumatol 66: 2124-2133.

248. Soloski MJ, Crowder LA, Lahey LJ, Wagner CA, Robinson WH, et al. (2014) Serum inflammatory mediators as markers of human Lyme disease activity. PLoS One 9: e93243.

249. Tang KS, Klempner MS, Wormser GP, Marques AR, Alaedini A (2015) Association of immune response to endothelial cell growth factor with early disseminated and late manifestations of Lyme disease but not posttreatment Lyme disease syndrome. Clin Infect Dis 61: 1703-1706.

250. Strle K, Sulka KB, Pianta A, Crowley JT, Arvikar SL, et al. (2017) T-Helper 17 Cell Cytokine Responses in Lyme Disease Correlate With Borrelia burgdorferi Antibodies During Early Infection and With Autoantibodies Late in the Illness in Patients With Antibiotic-Refractory Lyme Arthritis. Clin Infect Dis 64: 930-938.

251. Bouquet J, Soloski MJ, Swei A, Cheadle C, Federman S, et al. (2016) Longitudinal Transcriptome Analysis Reveals a Sustained Differential Gene Expression Signature in Patients Treated for Acute Lyme Disease. MBio 7: e00100-00116.

252. Morris G, Berk M, Walder K, Maes M (2016) The Putative Role of Viruses, Bacteria, and Chronic Fungal Biotoxin Exposure in the Genesis of Intractable Fatigue Accompanied by Cognitive and Physical Disability. Mol Neurobiol 53: 2550-2571.

253. Wyller VB, Eriksen HR, Malterud K (2009) Can sustained arousal explain the Chronic Fatigue Syndrome? Behav Brain Funct 5: 10.

254. Katz BZ, Jason LA (2013) Chronic fatigue syndrome following infections in adolescents. Curr Opin Pediatr 25: 95-102.

255. Morroy G, Keijmel SP, Delsing CE, Bleijenberg G, Langendam M, et al. (2016) Fatigue following Acute Q-Fever: A Systematic Literature Review. PLoS One 11: e0155884.

256. Hickie I, Davenport T, Wakefield D, Vollmer-Conna U, Cameron B, et al. (2006) Post-infective and chronic fatigue syndromes precipitated by viral and non-viral pathogens: prospective cohort study. BMJ 333: 575.

257. Hjetland R, Reiso H, Ihlebaek C, Nilsen RM, Grude N, et al. (2015) Subjective health complaints are not associated with tick bites or antibodies to Borrelia burgdorferi sensu lato in blood donors in western Norway: a cross-sectional study. BMC Public Health 15: 657.

258. Markowicz M, Kivaranovic D, Stanek G (2015) Testing patients with non-specific symptoms for antibodies against Borrelia burgdorferi sensu lato does not provide useful clinical information about their aetiology. Clin Microbiol Infect 21: 1098-1103.

259. Aguero-Rosenfeld ME, Wormser GP (2015) Lyme disease: diagnostic issues and controversies. Expert Rev Mol Diagn 15: 1-4.

260. Sinski E, Welc-Faleciak R, Zajkowska J (2016) Borrelia miyamotoi: A human tick-borne relapsing fever spirochete in Europe and its potential impact on public health. Adv Med Sci 61: 255-260.

261. Estrada-Pena A (2015) Ticks as vectors: taxonomy, biology and ecology. Rev Sci Tech 34: 53-65.

262. Rebaudet S, Parola P (2006) Epidemiology of relapsing fever borreliosis in Europe. FEMS Immunol Med Microbiol 48: 11-15.

263. Platonov AE, Karan LS, Kolyasnikova NM, Makhneva NA, Toporkova MG, et al. (2011) Humans infected with relapsing fever spirochete Borrelia miyamotoi, Russia. Emerg Infect Dis 17: 1816-1823.

264. Jungnick S, Margos G, Rieger M, Dzaferovic E, Bent SJ, et al. (2015) Borrelia burgdorferi sensu stricto and Borrelia afzelii: Population structure and differential pathogenicity. Int J Med Microbiol 305: 673-681.

265. Kurtenbach K, Hanincova K, Tsao JI, Margos G, Fish D, et al. (2006) Fundamental processes in the evolutionary ecology of Lyme borreliosis. Nat Rev Microbiol 4: 660-669.

266. Balmelli T, Piffaretti JC (1995) Association between different clinical manifestations of Lyme disease and different species of Borrelia burgdorferi sensu lato. Res Microbiol 146: 329-340.

267. Singh SK, Girschick HJ (2004) Lyme borreliosis: from infection to autoimmunity. Clin Microbiol Infect 10: 598-614.

268. Brisson D, Drecktrah D, Eggers CH, Samuels DS (2012) Genetics of Borrelia burgdorferi. Annu Rev Genet 46: 515-536.

269. Kenedy MR, Lenhart TR, Akins DR (2012) The role of Borrelia burgdorferi outer surface proteins. FEMS Immunol Med Microbiol 66: 1-19.

270. Wilske B, Preac-Mursic V, Jauris S, Hofmann A, Pradel I, et al. (1993) Immunological and molecular polymorphisms of OspC, an immunodominant major outer surface protein of Borrelia burgdorferi. Infect Immun 61: 2182-2191.

271. Probert WS, Crawford M, Cadiz RB, LeFebvre RB (1997) Immunization with outer surface protein (Osp) A, but not OspC, provides cross-protection of mice challenged with North American isolates of Borrelia burgdorferi. J Infect Dis 175: 400-405.

272. Seinost G, Dykhuizen DE, Dattwyler RJ, Golde WT, Dunn JJ, et al. (1999) Four clones of Borrelia burgdorferi sensu stricto cause invasive infection in humans. Infect Immun 67: 3518-3524.

273. Baranton G, Seinost G, Theodore G, Postic D, Dykhuizen D (2001) Distinct levels of genetic diversity of Borrelia burgdorferi are associated with different aspects of pathogenicity. Res Microbiol 152: 149-156.

274. Norris SJ (2014) vls Antigenic Variation Systems of Lyme Disease Borrelia: Eluding Host Immunity through both Random, Segmental Gene Conversion and Framework Heterogeneity. Microbiol Spectr 2.

275. Zhang JR, Norris SJ (1998) Genetic variation of the Borrelia burgdorferi gene vlsE involves cassette-specific, segmental gene conversion. Infect Immun 66: 3698-3704.

276. Liang FT, Alvarez AL, Gu Y, Nowling JM, Ramamoorthy R, et al. (1999) An immunodominant conserved region within the variable domain of VlsE, the variable surface antigen of Borrelia burgdorferi. J Immunol 163: 5566-5573.

277. Embers ME, Jacobs MB, Johnson BJ, Philipp MT (2007) Dominant epitopes of the C6 diagnostic peptide of Borrelia burgdorferi are largely inaccessible to antibody on the parent VlsE molecule. Clin Vaccine Immunol 14: 931-936.

278. Liang FT, Philipp MT (2000) Epitope mapping of the immunodominant invariable region of Borrelia burgdorferi VlsE in three host species. Infect Immun 68: 2349-2352.

279. Krupka I, Knauer J, Lorentzen L, O'Connor TP, Saucier J, et al. (2009) Borrelia burgdorferi sensu lato species in Europe induce diverse immune responses against C6 peptides in infected mice. Clin Vaccine Immunol 16: 1546-1562.

280. Iva Christova IT, Teodora Gladnishka, Evgenia Taseva, Vladislava Ivanova, Dobrinka Rusimova (2013) C6 Peptides from Borrelia burgdorferi sensu stricto, Borrelia afzelii and Borrelia garinii as Antigens for serological Diagnosis of Lyme Borreliosis. Biotechnol & Biotechnol Eq 27: 3540-3542.

281. Gomes-Solecki MJ, Meirelles L, Glass J, Dattwyler RJ (2007) Epitope length, genospecies dependency, and serum panel effect in the IR6 enzyme-linked immunosorbent assay for detection of antibodies to Borrelia burgdorferi. Clin Vaccine Immunol 14: 875-879.

282. Sillanpaa H, Lahdenne P, Sarvas H, Arnez M, Steere A, et al. (2007) Immune responses to borrelial VlsE IR6 peptide variants. Int J Med Microbiol 297: 45-52.

283. McDowell JV, Sung SY, Hu LT, Marconi RT (2002) Evidence that the variable regions of the central domain of VlsE are antigenic during infection with lyme disease spirochetes. Infect Immun 70: 4196-4203.

284. Jacek E, Tang KS, Komorowski L, Ajamian M, Probst C, et al. (2016) Epitope-Specific Evolution of Human B Cell Responses to Borrelia burgdorferi VlsE Protein from Early to Late Stages of Lyme Disease. J Immunol 196: 1036-1043.

285. Satz N (2010) Klinik der Lyme-Borreliose: Verlag Hans Huber, Hogrefe AG, Bern. 217 p.

286. Poljak A, Comstedt P, Hanner M, Schuler W, Meinke A, et al. (2012) Identification and characterization of Borrelia antigens as potential vaccine candidates against Lyme borreliosis. Vaccine 30: 4398-4406.

287. Menten-Dedoyart C, Couvreur B, Thellin O, Drion PV, Herry M, et al. (2008) Influence of the Ixodes ricinus tick blood-feeding on the antigen-specific antibody response in vivo. Vaccine 26: 6956-6964.

288. Menten-Dedoyart C, Couvreur B, Jolois O, Van Lerberghe PB, Duwez L, et al. (2011) Kinetic study of the antibody response during the blood meal of Ixodes ricinus: implication on plasma cell maturation in vivo and for anti-Ixodes vaccination. Vaccine 29: 2044-2050.

289. Hastey CJ, Elsner RA, Barthold SW, Baumgarth N (2012) Delays and diversions mark the development of B cell responses to Borrelia burgdorferi infection. J Immunol 188: 5612-5622.

290. Tunev SS, Hastey CJ, Hodzic E, Feng S, Barthold SW, et al. (2011) Lymphoadenopathy during lyme borreliosis is caused by spirochete migration-induced specific B cell activation. PLoS Pathog 7: e1002066.

291. Elsner RA, Hastey CJ, Baumgarth N (2015) CD4+ T cells promote antibody production but not sustained affinity maturation during Borrelia burgdorferi infection. Infect Immun 83: 48-56.

292. Elsner RA, Hastey CJ, Olsen KJ, Baumgarth N (2015) Suppression of Long-Lived Humoral Immunity Following Borrelia burgdorferi Infection. PLoS Pathog 11: e1004976.

293. Hastey CJ, Ochoa J, Olsen KJ, Barthold SW, Baumgarth N (2014) MyD88- and TRIF-independent induction of type I interferon drives naive B cell accumulation but not loss of lymph node architecture in Lyme disease. Infect Immun 82: 1548-1558.

294. Hannier S, Liversidge J, Sternberg JM, Bowman AS (2004) Characterization of the B-cell inhibitory protein factor in Ixodes ricinus tick saliva: a potential role in enhanced Borrelia burgdoferi transmission. Immunology 113: 401-408.

295. Kalish RA, McHugh G, Granquist J, Shea B, Ruthazer R, et al. (2001) Persistence of immunoglobulin M or immunoglobulin G antibody responses to Borrelia burgdorferi 10-20 years after active Lyme disease. Clin Infect Dis 33: 780-785.

296. Crotty S, Johnston RJ, Schoenberger SP (2010) Effectors and memories: Bcl-6 and Blimp-1 in T and B lymphocyte differentiation. Nat Immunol 11: 114-120.

297. Krupna-Gaylord MA, Liveris D, Love AC, Wormser GP, Schwartz I, et al. (2014) Induction of type I and type III interferons by Borrelia burgdorferi correlates with pathogenesis and requires linear plasmid 36. PLoS One 9: e100174.

298. Schoenfeld R, Araneo B, Ma Y, Yang LM, Weis JJ (1992) Demonstration of a B-lymphocyte mitogen produced by the Lyme disease pathogen, Borrelia burgdorferi. Infect Immun 60: 455-464.

299. Dickinson GS, Alugupalli KR (2012) Deciphering the role of Toll-like receptors in humoral responses to Borreliae. Front Biosci (Schol Ed) 4: 699-712.

300. Carroll MC (2000) The role of complement in B cell activation and tolerance. Adv Immunol 74: 61-88.

301. Kraiczy P (2016) Hide and Seek: How Lyme Disease Spirochetes Overcome Complement Attack. Front Immunol 7: 385.

302. Randolph SE, Gern L, Nuttall PA (1996) Co-feeding ticks: Epidemiological significance for tick-borne pathogen transmission. Parasitol Today 12: 472-479.

303. Tracy KE, Baumgarth N (2017) Borrelia burgdorferi Manipulates Innate and Adaptive Immunity to Establish Persistence in Rodent Reservoir Hosts. Front Immunol 8: 116.

304. Glatz M, Golestani M, Kerl H, Mullegger RR (2006) Clinical relevance of different IgG and IgM serum antibody responses to Borrelia burgdorferi after antibiotic therapy for erythema migrans: long-term follow-up study of 113 patients. Arch Dermatol 142: 862-868.

305. Lomholt H, Lebech AM, Hansen K, Brandrup F, Halkier-Sorensen L (2000) Long-term serological follow-up of patients treated for chronic cutaneous borreliosis or culture-positive erythema migrans. Acta Derm Venereol 80: 362-366.

306. Rebman AW, Crowder LA, Kirkpatrick A, Aucott JN (2015) Characteristics of seroconversion and implications for diagnosis of post-treatment Lyme disease syndrome: acute and convalescent serology among a prospective cohort of early Lyme disease patients. Clin Rheumatol 34: 585-589.

307. Ang CW, Notermans DW, Hommes M, Simoons-Smit AM, Herremans T (2011) Large differences between test strategies for the detection of anti-Borrelia antibodies are revealed by comparing eight ELISAs and five immunoblots. Eur J Clin Microbiol Infect Dis 30: 1027-1032.

308. Aguero-Rosenfeld ME, Nowakowski J, Bittker S, Cooper D, Nadelman RB, et al. (1996) Evolution of the serologic response to Borrelia burgdorferi in treated patients with culture-confirmed erythema migrans. J Clin Microbiol 34: 1-9.

309. Hammers-Berggren S, Lebech AM, Karlsson M, Svenungsson B, Hansen K, et al. (1994) Serological follow-up after treatment of patients with erythema migrans and neuroborreliosis. J Clin Microbiol 32: 1519-1525.

310. Steere AC, Strle F, Wormser GP, Hu LT, Branda JA, et al. (2016) Lyme borreliosis. Nat Rev Dis Primers 2: 16090.

311. Krause PJ, Foley DT, Burke GS, Christianson D, Closter L, et al. (2006) Reinfection and relapse in early Lyme disease. Am J Trop Med Hyg 75: 1090-1094.

312. Fikrig E, Bockenstedt LK, Barthold SW, Chen M, Tao H, et al. (1994) Sera from patients with chronic Lyme disease protect mice from Lyme borreliosis. J Infect Dis 169: 568-574.

313. Pavia CS, Wormser GP, Norman GL (1997) Activity of sera from patients with Lyme disease against Borrelia burgdorferi. Clin Infect Dis 25 Suppl 1: S25-30.

314. Comstedt P, Hanner M, Schuler W, Meinke A, Lundberg U (2014) Design and development of a novel vaccine for protection against Lyme borreliosis. PLoS One 9: e113294.

315. Wressnigg N, Pollabauer EM, Aichinger G, Portsmouth D, Low-Baselli A, et al. (2013) Safety and immunogenicity of a novel multivalent OspA vaccine against Lyme borreliosis in healthy adults: a double-blind, randomised, dose-escalation phase 1/2 trial. Lancet Infect Dis 13: 680-689.

316. Izac JR, Oliver LD, Jr., Earnhart CG, Marconi RT (2017) Identification of a defined linear epitope in the OspA protein of the Lyme disease spirochetes that elicits bactericidal antibody responses: Implications for vaccine development. Vaccine 35: 3178-3185.

317. Janols H, Bredberg A, Thuvesson I, Janciauskiene S, Grip O, et al. (2010) Lymphocyte and monocyte flow cytometry immunophenotyping as a diagnostic tool in uncharacteristic inflammatory disorders. BMC Infect Dis 10: 205.

318. Salazar JC, Pope CD, Sellati TJ, Feder HM, Jr., Kiely TG, et al. (2003) Coevolution of markers of innate and adaptive immunity in skin and peripheral blood of patients with erythema migrans. J Immunol 171: 2660-2670.

319. Baig S, Olsson T, Link H (1989) Predominance of Borrelia burgdorferi specific B cells in cerebrospinal fluid in neuroborreliosis. Lancet 2: 71-74.

320. Lakos A, Ferenczi E, Komoly S, Granstrom M (2005) Different B-cell populations are responsible for the peripheral and intrathecal antibody production in neuroborreliosis. Int Immunol 17: 1631-1637.

321. Steere AC, Duray PH, Butcher EC (1988) Spirochetal antigens and lymphoid cell surface markers in Lyme synovitis. Comparison with rheumatoid synovium and tonsillar lymphoid tissue. Arthritis Rheum 31: 487-495.

322. Ghosh S, Steere AC, Stollar BD, Huber BT (2005) In situ diversification of the antibody repertoire in chronic Lyme arthritis synovium. J Immunol 174: 2860-2869.

323. Sigal LH, Steere AC, Freeman DH, Dwyer JM (1986) Proliferative responses of mononuclear cells in Lyme disease. Reactivity to Borrelia burgdorferi antigens is greater in joint fluid than in blood. Arthritis Rheum 29: 761-769.

324. Jones GW, Jones SA (2016) Ectopic lymphoid follicles: inducible centres for generating antigen-specific immune responses within tissues. Immunology 147: 141-151.

325. Strle K, Sulka KB, Pianta A, Crowley JT, Arvikar SL, et al. (2017) TH17 Cytokine Responses in Lyme Disease Correlate with Borreliaburgdorferi Antibodies During Early Infection in Patients with Erythema Migrans and with Autoantibodies Late in the Illness in Patients with Antibiotic-Refractory Lyme Arthritis. Clin Infect Dis.

326. Truck J, Ramasamy MN, Galson JD, Rance R, Parkhill J, et al. (2015) Identification of antigen-specific B cell receptor sequences using public repertoire analysis. J Immunol 194: 252-261.

327. Wormser GP, Schriefer M, Aguero-Rosenfeld ME, Levin A, Steere AC, et al. (2013) Single-tier testing with the C6 peptide ELISA kit compared with two-tier testing for Lyme disease. Diagn Microbiol Infect Dis 75: 9-15.

328. Racanelli V, Brunetti C, De Re V, Caggiari L, De Zorzi M, et al. (2011) Antibody V(h) repertoire differences between resolving and chronically evolving hepatitis C virus infections. PLoS One 6: e25606.

329. Michel H, Wilske B, Hettche G, Gottner G, Heimerl C, et al. (2004) An ospA-polymerase chain reaction/restriction fragment length polymorphism-based method for sensitive detection and reliable differentiation of all European Borrelia burgdorferi sensu lato species and OspA types. Med Microbiol Immunol 193: 219-226.

330. Ishikura M, Ando S, Shinagawa Y, Matsuura K, Hasegawa S, et al. (2003) Phylogenetic analysis of spotted fever group rickettsiae based on gltA, 17-kDa, and rOmpA genes amplified by nested PCR from ticks in Japan. Microbiol Immunol 47: 823-832.

331. Wang X, Stollar BD (2000) Human immunoglobulin variable region gene analysis by single cell RT-PCR. J Immunol Methods 244: 217-225.

332. Ogunniyi AO, Thomas BA, Politano TJ, Varadarajan N, Landais E, et al. (2014) Profiling human antibody responses by integrated single-cell analysis. Vaccine 32: 2866-2873.

333. Tiller T, Meffre E, Yurasov S, Tsuiji M, Nussenzweig MC, et al. (2008) Efficient generation of monoclonal antibodies from single human B cells by single cell RT-PCR and expression vector cloning. J Immunol Methods 329: 112-124.

334. Smith K, Garman L, Wrammert J, Zheng NY, Capra JD, et al. (2009) Rapid generation of fully human monoclonal antibodies specific to a vaccinating antigen. Nat Protoc 4: 372-384.

335. van Dongen JJ, Langerak AW, Bruggemann M, Evans PA, Hummel M, et al. (2003) Design and standardization of PCR primers and protocols for detection of clonal immunoglobulin and T-cell receptor gene recombinations in suspect lymphoproliferations: report of the BIOMED-2 Concerted Action BMH4-CT98-3936. Leukemia 17: 2257-2317.

336. Fu L, Niu B, Zhu Z, Wu S, Li W (2012) CD-HIT: accelerated for clustering the next-generation sequencing data. Bioinformatics 28: 3150-3152.

337. Li W, Godzik A (2006) Cd-hit: a fast program for clustering and comparing large sets of protein or nucleotide sequences. Bioinformatics 22: 1658-1659.

338. Gupta NT, Vander Heiden JA, Uduman M, Gadala-Maria D, Yaari G, et al. (2015) Change-O: a toolkit for analyzing large-scale B cell immunoglobulin repertoire sequencing data. Bioinformatics 31: 3356-3358.

339. Ye J, Ma N, Madden TL, Ostell JM (2013) IgBLAST: an immunoglobulin variable domain sequence analysis tool. Nucleic Acids Res 41: W34-40.

340. Alamyar E, Duroux P, Lefranc MP, Giudicelli V (2012) IMGT((R)) 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. Methods Mol Biol 882: 569-604.

341. Brochet X, Lefranc MP, Giudicelli V (2008) IMGT/V-QUEST: the highly customized and integrated system for IG and TR standardized V-J and V-D-J sequence analysis. Nucleic Acids Res 36: W503-508.

342. Loytynoja A, Vilella AJ, Goldman N (2012) Accurate extension of multiple sequence alignments using a phylogeny-aware graph algorithm. Bioinformatics 28: 1684-1691.

343. Dereeper A, Audic S, Claverie JM, Blanc G (2010) BLAST-EXPLORER helps you building datasets for phylogenetic analysis. BMC Evol Biol 10: 8.

344. Dereeper A, Guignon V, Blanc G, Audic S, Buffet S, et al. (2008) Phylogeny.fr: robust phylogenetic analysis for the non-specialist. Nucleic Acids Res 36: W465-469.

345. Edgar RC (2004) MUSCLE: multiple sequence alignment with high accuracy and high throughput. Nucleic Acids Res 32: 1792-1797.

346. Castresana J (2000) Selection of conserved blocks from multiple alignments for their use in phylogenetic analysis. Mol Biol Evol 17: 540-552.

347. Stucky BJ (2012) SeqTrace: a graphical tool for rapidly processing DNA sequencing chromatograms. J Biomol Tech 23: 90-93.

348. Shugay M, Bagaev DV, Turchaninova MA, Bolotin DA, Britanova OV, et al. (2015) VDJtools: Unifying Post-analysis of T Cell Receptor Repertoires. PLoS Comput Biol 11: e1004503.

349. Reye AL, Hubschen JM, Sausy A, Muller CP (2010) Prevalence and seasonality of tick-borne pathogens in questing Ixodes ricinus ticks from Luxembourg. Appl Environ Microbiol 76: 2923-2931.

350. Gray J, Stanek G, Kundi M, Kocianova E (2005) Dimensions of engorging Ixodes ricinus as a measure of feeding duration. Int J Med Microbiol 295: 567-572.

351. Meiners T, Hammer B, Gobel UB, Kahl O (2006) Determining the tick scutal index allows assessment of tick feeding duration and estimation of infection risk with Borrelia burgdorferi sensu lato in a person bitten by an Ixodes ricinus nymph. Int J Med Microbiol 296 Suppl 40: 103-107.

352. Franz B, May KF, Jr., Dranoff G, Wucherpfennig K (2011) Ex vivo characterization and isolation of rare memory B cells with antigen tetramers. Blood 118: 348-357.

353. Ouisse LH, Gautreau-Rolland L, Devilder MC, Osborn M, Moyon M, et al. (2017) Antigen-specific single B cell sorting and expression-cloning from immunoglobulin humanized rats: a rapid and versatile method for the generation of high affinity and discriminative human monoclonal antibodies. BMC Biotechnol 17: 3.

354. Li S, Lefranc MP, Miles JJ, Alamyar E, Giudicelli V, et al. (2013) IMGT/HighV QUEST paradigm for T cell receptor IMGT clonotype diversity and next generation repertoire immunoprofiling. Nat Commun 4: 2333.

355. Huang J, Doria-Rose NA, Longo NS, Laub L, Lin CL, et al. (2013) Isolation of human monoclonal antibodies from peripheral blood B cells. Nat Protoc 8: 1907-1915.

356. Palanichamy A, Apeltsin L, Kuo TC, Sirota M, Wang S, et al. (2014) Immunoglobulin class-switched B cells form an active immune axis between CNS and periphery in multiple sclerosis. Sci Transl Med 6: 248ra106.

357. Romme Christensen J, Bornsen L, Ratzer R, Piehl F, Khademi M, et al. (2013) Systemic inflammation in progressive multiple sclerosis involves follicular T-helper, Th17- and activated B-cells and correlates with progression. PLoS One 8: e57820.

358. Rivas JR, Ireland SJ, Chkheidze R, Rounds WH, Lim J, et al. (2017) Peripheral VH4+ plasmablasts demonstrate autoreactive B cell expansion toward brain antigens in early multiple sclerosis patients. Acta Neuropathol 133: 43-60.

359. Claes N, Fraussen J, Stinissen P, Hupperts R, Somers V (2015) B Cells Are Multifunctional Players in Multiple Sclerosis Pathogenesis: Insights from Therapeutic Interventions. Front Immunol 6: 642.

360. Habib J, Deng J, Lava N, Tyor W, Galipeau J (2015) Blood B Cell and Regulatory Subset Content in Multiple Sclerosis Patients. J Mult Scler (Foster City) 2.

361. Agathangelidis A, Darzentas N, Hadzidimitriou A, Brochet X, Murray F, et al. (2012) Stereotyped B-cell receptors in one-third of chronic lymphocytic leukemia: a molecular classification with implications for targeted therapies. Blood 119: 4467-4475.

362. Sutton LA, Young E, Baliakas P, Hadzidimitriou A, Moysiadis T, et al. (2016) Different spectra of recurrent gene mutations in subsets of chronic lymphocytic leukemia harboring stereotyped B-cell receptors. Haematologica 101: 959-967.

363. Kyu SY, Kobie J, Yang H, Zand MS, Topham DJ, et al. (2009) Frequencies of human influenza-specific antibody secreting cells or plasmablasts post vaccination from fresh and frozen peripheral blood mononuclear cells. J Immunol Methods 340: 42-47.

364. Best K, Oakes T, Heather JM, Shawe-Taylor J, Chain B (2015) Computational analysis of stochastic heterogeneity in PCR amplification efficiency revealed by single molecule barcoding. Sci Rep 5: 14629.

365. Mamedov IZ, Britanova OV, Zvyagin IV, Turchaninova MA, Bolotin DA, et al. (2013) Preparing unbiased T-cell receptor and antibody cDNA libraries for the deep next generation sequencing profiling. Front Immunol 4: 456.

366. Carlson CS, Emerson RO, Sherwood AM, Desmarais C, Chung MW, et al. (2013) Using synthetic templates to design an unbiased multiplex PCR assay. Nat Commun 4: 2680.

367. Bashford-Rogers RJ, Palser AL, Huntly BJ, Rance R, Vassiliou GS, et al. (2013) Network properties derived from deep sequencing of human B-cell receptor repertoires delineate B-cell populations. Genome Res 23: 1874-1884.

368. Schutze T, Rubelt F, Repkow J, Greiner N, Erdmann VA, et al. (2011) A streamlined protocol for emulsion polymerase chain reaction and subsequent purification. Anal Biochem 410: 155-157.

369. Bragg LM, Stone G, Butler MK, Hugenholtz P, Tyson GW (2013) Shining a light on dark sequencing: characterising errors in Ion Torrent PGM data. PLoS Comput Biol 9: e1003031.

370. Laehnemann D, Borkhardt A, McHardy AC (2016) Denoising DNA deep sequencing data-high-throughput sequencing errors and their correction. Brief Bioinform 17: 154-179.

371. Brodin J, Hedskog C, Heddini A, Benard E, Neher RA, et al. (2015) Challenges with using primer IDs to improve accuracy of next generation sequencing. PLoS One 10: e0119123.

372. Bashford-Rogers RJ, Palser AL, Idris SF, Carter L, Epstein M, et al. (2014) Capturing needles in haystacks: a comparison of B-cell receptor sequencing methods. BMC Immunol 15: 29.

373. Laserson U, Vigneault F, Gadala-Maria D, Yaari G, Uduman M, et al. (2014) High-resolution antibody dynamics of vaccine-induced immune responses. Proc Natl Acad Sci U S A 111: 4928-4933.

374. Jiang N, He J, Weinstein JA, Penland L, Sasaki S, et al. (2013) Lineage structure of the human antibody repertoire in response to influenza vaccination. Sci Transl Med 5: 171ra119.

375. Traggiai E (2012) Immortalization of human B cells: analysis of B cell repertoire and production of human monoclonal antibodies. Methods Mol Biol 901: 161-170.

376. Traggiai E, Becker S, Subbarao K, Kolesnikova L, Uematsu Y, et al. (2004) An efficient method to make human monoclonal antibodies from memory B cells: potent neutralization of SARS coronavirus. Nat Med 10: 871-875.

377. Kwakkenbos MJ, Diehl SA, Yasuda E, Bakker AQ, van Geelen CM, et al. (2010) Generation of stable monoclonal antibody-producing B cell receptor-positive human memory B cells by genetic programming. Nat Med 16: 123-128.

378. Weitkamp JH, Kallewaard N, Kusuhara K, Feigelstock D, Feng N, et al. (2003) Generation of recombinant human monoclonal antibodies to rotavirus from single antigen-specific B cells selected with fluorescent virus-like particles. J Immunol Methods 275: 223-237.

379. Ghosh S, Huber BT (2007) Clonal diversification in OspA-specific antibodies from peripheral circulation of a chronic Lyme arthritis patient. J Immunol Methods 321: 121-134.

380. Lingwood D, McTamney PM, Yassine HM, Whittle JR, Guo X, et al. (2012) Structural and genetic basis for development of broadly neutralizing influenza antibodies. Nature 489: 566-570.

381. Zhu J, Wu X, Zhang B, McKee K, O'Dell S, et al. (2013) De novo identification of VRC01 class HIV-1-neutralizing antibodies by next-generation sequencing of B-cell transcripts. Proc Natl Acad Sci U S A 110: E4088-4097.

382. Singh SK, Girschick HJ (2004) Molecular survival strategies of the Lyme disease spirochete Borrelia burgdorferi. Lancet Infect Dis 4: 575-583.

383. Burke G, Wikel SK, Spielman A, Telford SR, McKay K, et al. (2005) Hypersensitivity to ticks and Lyme disease risk. Emerg Infect Dis 11: 36-41.

384. Tsioris K, Gupta NT, Ogunniyi AO, Zimnisky RM, Qian F, et al. (2015) Neutralizing antibodies against West Nile virus identified directly from human B cells by single-cell analysis and next generation sequencing. Integr Biol (Camb) 7: 1587-1597.

385. Shahaf G, Barak M, Zuckerman NS, Swerdlin N, Gorfine M, et al. (2008) Antigen-driven selection in germinal centers as reflected by the shape characteristics of immunoglobulin gene lineage trees: a large-scale simulation study. J Theor Biol 255: 210-222.

386. https://www.cdc.gov/tetanus/about/causes-transmission.html.

387. Hainz U, Jenewein B, Asch E, Pfeiffer KP, Berger P, et al. (2005) Insufficient protection for healthy elderly adults by tetanus and TBE vaccines. Vaccine 23: 3232-3235.

388. Thwaites CL, Loan HT (2015) Eradication of tetanus. Br Med Bull 116: 69-77.

389. D'Arco C, Dattwyler RJ, Arnaboldi PM (2017) Borrelia burgdorferi-specific IgA in Lyme Disease. EBioMedicine 19: 91-97.

390. Buisman AM, de Rond CG, Ozturk K, Ten Hulscher HI, van Binnendijk RS (2009) Long-term presence of memory B-cells specific for different vaccine components. Vaccine 28: 179-186.

8 ANNEX

8.1 Questionnaires

The questionnaires for high risk individuals have been setup by Anna L. Reye and the ones for the acute patients were adapted together with Josiane Kirpach. The ok from the ethics committee and data protection agency have already been obtained prior to the start of Josiane's PhD.

8.1.1 Fragebogen Risikogruppe „Neue Biomarker für Borreliose"

1. Allgemeine Angaben

☐ männlich ☐ weiblich Geburtsjahr:.....

Wohnort oder PLZ:..... Beruf:.....

Haben Sie bereits vorher an der Zeckenstudie der Abteilung für Immunologie des CRP-Santé teilgenommen?

☐ Ja ☐ Nein

2. Durchschnittliche Aufenthaltsdauer im Freien von März bis Oktober (Zeckenaktivitätsperiode)

Berufsbedingt:.....(Stunden pro Tag) Privat: .....(Stunden pro Tag)

3. Wie viele Zecken sammeln Sie jährlich von sich ab? (inkl. derjenigen, die sich noch nicht festgebissen hatten)

☐ keine Zecke ☐ 1-5 Zecken ☐ 6-10 Zecken ☐ 11-20 Zecken ☐ 21-50 Zecken ☐ 51-100 Zecken ☐ >100 Zecken

4. Wie viele Zeckenstiche stellen Sie jährlich an sich fest?

☐ keine Zecke ☐ 1-5 Zecken ☐ 6-10 Zecken ☐ 11-20 Zecken ☐ 21-50 Zecken ☐ 51-100 Zecken ☐ >100 Zecken

Wann stellten Sie den letzten Zeckenstich bei sich fest?.....

5. Ergreifen Sie regelmäßig Präventionsmassnahmen? (Mehrfachnennungen möglich)

☐ Zeckenabwehrmittel ☐ Körperinspektion nach Aufenthalt in Zeckengebieten
☐ Frühes Entfernen festgebissener Zecken ☐ Kleidung als Schutz
☐ Andere (genaue Angaben).....

6. Beobachten Sie die Einstichstelle nach Entfernung der Zecke?

☐ Ja, für ca. ..... Tage ☐ Nein

7. Wurde bei Ihnen bereits Lyme Borreliose diagnostiziert? Wenn ja, in welcher Manifestation?

☐ Ja ☐ Nein
☐ Erythema migrans (Wanderröte) ☐ Neuroborreliose ☐ Lyme Arthritis
☐ Acrodermatitis chronica atrophicans ☐ andere: .....

Wann wurde die Diagnose erstellt? .....

8. Wie erfolgreich wurde die Lyme Borreliose behandelt?

☐ Vollständige Heilung
☐ Keine vollständige Heilung, verbliebene Symptome sind.....
☐ Therapie schlug fehl, verbliebene Symptome sind.....

Weitere Kommentare: .....

8.1.2 Questionnaire groupe à risque „Nouveaux biomarqueurs pour la borréliose“

1. Données générales

☐ masculin ☐ féminin Année de naissance : .....

Domicile ou code postal : ..... Profession : .....

Avez-vous déjà participé auparavant à l'étude sur les tiques du Département d'Immunologie du CRP-Santé?

☐ Oui ☐ Non

2. Durée moyenne de séjour à l'air libre pendant la période de mars – octobre (période d'activité des tiques)

Raisons Professionnelles : ..... (Heures/jour) et/ou Privées : ..... (Heures/jour)

3. Sur une période d'un an, combien de tiques avez-vous détecté sur votre corps ? (y compris les non fixées)

☐ Aucune tique ☐ 1-5 tiques ☐ 6-10 tiques ☐ 11-20 tiques ☐ 21-50 tiques ☐ 51-100 tiques
☐ >100 tiques

4. En moyenne, combien de piqûres de tiques avez-vous eu par an ?

☐ Aucune piqûre ☐ 1-5 piqûres ☐ 6-10 piqûres ☐ 11-20 piqûres ☐ 21-50 piqûres ☐ 51-100 piqûres
☐ >100 piqûres

Quand avez-vous remarqué votre dernière piqûre de tique ? .....

5. Est-ce que vous adoptez régulièrement des mesures préventives ? (plusieurs réponses sont possibles)

☐ Utilisation de répulsifs à tiques ☐ Inspection corporelle après un séjour en plein air
☐ Enlèvement précoce des tiques fixées ☐ Port de vêtements adaptés (couvrants)
☐ Autres (indications précises) .....

6. Observez-vous l'endroit de piqûre fait par la tique ?

☐ Oui, pendant environ ..... jours ☐ Non

7. Avez-vous été diagnostiqué comme ayant la maladie de Lyme ? Si oui, sur la base de quelle(s) manisfestation(s)?

☐ Oui ☐ Non

☐ Érythème migrant ☐ Neuroborréliose

☐ Arthrite de Lyme ☐ Acrodermatite chronique atrophiante (ACA)

☐ Autres.....

Quand avez-vous été diagnostiqué ? .....

8. Avec quel succès a-t-on traité la borréliose de Lyme ?

☐ Guérison complète

☐ Guérison incomplète, les symptômes restants sont : .....

☐ La thérapie a échoué, les symptômes restants sont : .....

Commentaires supplémentaires : .....

8.1.3 Fragebogen bei Erstvorstellung

1. Unter welchen Symptomen leiden Sie zurzeit? (Mehrfachnennungen möglich)		
☐ Keine Symptome	**Neuroborreliose:**	**Hautveränderungen:**
Allgemeine Symptome: ☐ Kopfschmerzen ☐ Grippe-ähnliche Symptome ☐ Nackensteifheit ☐ Gleichgewichtsstörungen ☐ Konzentrationsstörungen ☐ Gedächtnisstörungen ☐ Aussergewöhnliche Müdigkeit ☐ Schlafstörungen ☐ Wesensveränderung, Depression	☐ Nervenschmerzen ☐ Entzündungen des Gehirns ☐ Lähmungserscheinungen ☐ Sensibilitätsstörungen **Arthritis:** ☐ Gelenkschmerzen Betroffene Gelenke: ..... ☐ Muskelschmerzen	☐ Wanderröte (Erythema migrans) ☐ Bläulich-rote Verfärbungen der Haut ☐ Schwellungen der Haut ☐ Zigarettenpapierartig gefältete Haut ☐ Ungewöhnlich dünne Haut ☐ Durchschimmernde Gefässe ☐ Knotenbildung der Haut **Andere Symptome:**

Sollten Sie an Hautveränderungen leiden, können Sie gerne ein Foto beilegen oder an unsere Emailadresse borreliose@crp-sante.lu schicken.

2. Seit wann beobachten Sie die oben genannten Beschwerden?

.....

3. Wie wurde die Diagnose erstellt?

☐ Klinische Symptome ☐ Serologie ☐ Andere Tests: .....

4. Serologisches Ergebnis

Borrelia burgd. IgG:	.....	Bor. burgd. Blot IgG:	.....
Borrelia burgd. IgG Index:	.....		
Borrelia burgd. IgM:	.....	Bor. burgd. Blot IgM:	.....
Borrelia burgd. IgM Index:	.....		

Wenn möglich legen Sie bitte jetzt oder beim nächsten Termin eine Kopie Ihres serologischen Befundes bei.

5. Therapie

Medikament/Antibiotikum: ..... Therapiebeginn: ..... Dauer der Therapie: .....

6. Durchschnittliche Aufenthaltsdauer im Freien von März bis Oktober (Zeckenaktivitätsperiode)

Berufsbedingt:.....(Stunden pro Tag) Privat: .....(Stunden pro Tag)

7. Bei welcher Beschäftigung haben Sie den meisten Kontakt zu Zecken?

.....

8. Wieviele Zecken sammeln Sie jährlich von sich ab? (inkl. noch nicht festgebissener Zecken)

☐ keine ☐ 1-5 Zecken ☐ 6-10 Zecken ☐ 11-20 Zecken ☐ 21-50 Zecken ☐ 51-100 Zecken ☐ >100 Zecken

9. Wieviele Zeckenstiche stellen Sie jährlich bei sich fest?

☐ keine ☐ 1-5 Zecken ☐ 6-10 Zecken ☐ 11-20 Zecken ☐ 21-50 Zecken ☐ 51-100 Zecken ☐ >100 Zecken

Wann stellten Sie den letzten Zeckenstich bei sich fest (ungefähr)?.....

10. Ergreifen Sie regelmäßig Präventionsmassnahmen? (Mehrfachnennungen möglich)

☐ Zeckenabwehrmittel
☐ Körperinspektion nach Aufenthalt in Zeckengebieten
☐ Frühes Entfernen festgebissener Zecken
☐ Kleidung als Schutz
☐ Andere (genaue Angaben).....

11. Beobachten Sie die Einstichstelle nach Entfernung der Zecke?

☐ Normalerweise ja, für ca. ..... Tage
☐ Nein

12. Hat Ihr Arzt bei Ihnen schon einmal eine Borreliose diagnostiziert?

.....
.....

8.1.4 Fragebogen für Folgetermine

1. Unter welchen Symptomen leiden Sie zurzeit? (Mehrfachnennungen möglich)

☐ Keine Symptome	**Neuroborreliose:** ☐ Nervenschmerzen ☐ Entzündungen des Gehirns ☐ Lähmungserscheinungen ☐ Sensibilitätsstörungen	**Hautveränderungen:** ☐ Wanderröte (Erythema migrans) ☐ Bläulich-rote Verfärbungen der Haut ☐ Schwellungen der Haut ☐ Zigarettenpapierartig gefältete Haut ☐ Ungewöhnlich dünne Haut ☐ Durchschimmernde Gefässe ☐ Knotenbildung der Haut
Allgemeine Symptome: ☐ Kopfschmerzen ☐ Grippe-ähnliche Symptome ☐ Nackensteifheit ☐ Gleichgewichtsstörungen ☐ Konzentrationsstörungen ☐ Gedächtnisstörungen ☐ Aussergewöhnliche Müdigkeit ☐ Schlafstörungen ☐ Wesensveränderung, Depression	**Arthritis:** ☐ Gelenkschmerzen Betroffene Gelenke: ..... ☐ Muskelschmerzen	**Andere Symptome:**

Sollten Sie an Hautveränderungen leiden, können Sie gerne ein Foto beilegen oder an unsere Emailadresse borreliose@crp-sante.lu schicken.

2. Seit wann treten die oben genannten Beschwerden auf?

.....

3. Therapie

Medikament/Antibiotikum: ..... Therapiebeginn: ..... Dauer der Therapie: .....

4. Weitere Kommentare

.....
.....
.....
.....

Bitte beantworten Sie auch die folgenden Fragen, die im Rahmen einer anderen Studie der Abteilung für Immunologie von Bedeutung sind:

5. Sind Sie schon einmal an Masern erkrankt?

☐ Ja und zwar ungefähr vor ..... Jahren
☐ Nein
☐ Weiss nicht

6. Wurden Sie gegen Masern geimpft?

☐ Ja und zwar ungefähr vor ..... Jahren
☐ Nein
☐ Weiss nicht

8.1.5 Questionnaire première séance

1. Quels sont les symptômes que vous présentez en ce moment? (Plusieurs réponses possibles)

☐ pas de symptômes	**Neuroborréliose:** ☐ Névralgie ☐ Inflammation du cerveau ☐ Paralysie ☐ Troubles sensoriels	**Manifestations sur la peau :** ☐ Érythème migrant ☐ Décoloration rouge-bleuâtre ☐ Enflure ☐ Apparence craquelée rappelant un papier de cigarette froissé ☐ Peau anormalement mince ☐ Vaisseaux sanguins transparaissant ☐ Nœuds sur la peau
Symptômes générales: ☐ Maux de tête ☐ Symptômes grippaux ☐ Raideur de la nuque ☐ Troubles d'équilibre ☐ Troubles de concentration ☐ Troubles de la mémoire ☐ Fatigue exceptionnelle ☐ Insomnie ☐ Changement de la personnalité, Dépression	**Arthrite:** ☐ Arthralgie Articulations affectées: ..... ☐ Douleurs musculaires	**Autres symptômes:**

Si vous souffrez de manifestations de la peau, vous pouvez joindre une photo ou nous l'envoyer à l'adresse e-mail borreliose@crp-sante.lu **.**

2. Depuis quand observez-vous les symptômes susmentionnés?

.....

3. Comment la borréliose a-t-elle été diagnostiquée?

☐ Symptômes cliniques ☐ Sérologie ☐ Autres tests: .....

4. Résultats sérologiques

Borrelia burgd. IgG: ..... Bor. burgd. blot IgG: .....
Borrelia burgd. IgG index: .....
Borrelia burgd. IgM: ..... Bor. burgd. blot IgM: .....
Borrelia burgd. IgM index: .....

Si possible, veuillez joindre une copie des tests sérologiques effectués ou les emmener lors du prochain rendez-vous.

5. Thérapie

Médicament/Antibiotique: ..... Début de la thérapie: ..... Durée de la thérapie: .....

6. Durée moyenne de séjour à l'extérieur pendant la période de mars - octobre (période d'activité des tiques)

Raisons Professionnelles: ..... (Heures/jour) et/ou Privées : ..... (Heures/jour)

7. Durant quelle(s) activité(s) avez-vous plus de risque d'expositions aux tiques?

.....

8. Sur une période d'un an, combien de tiques avez-vous détecté sur votre corps? (y compris les non fixées)

☐ Aucune ☐ 1-5 tiques ☐ 6-10 tiques ☐ 11-20 tiques ☐ 21-50 tiques ☐ 51-100 tiques ☐ >100 tiques

9. En moyenne, combien de piqûres de tiques complez-vous annuellement?

☐ Aucune ☐ 1-5 tiques ☐ 6-10 tiques ☐ 11-20 tiques ☐ 21-50 tiques ☐ 51-100 tiques ☐ >100 tiques

Quand avez-vous remarqué votre dernière piqûre de tique (à peu près) ? .....

10. Est-ce que vous adoptez régulièrement des mesures préventives? (plusieurs réponses sont possibles)

☐ Utilisation de répulsifs à tiques ☐ Inspection corporelle après un séjour en plein air

☐ Enlèvement précoce des tiques fixées ☐ Port de vêtements adaptés (couvrants)

☐ Autres (indications précises) .....

11. Observez-vous l'endroit de piqûre après enlèvement de la tique?

☐ Normalement oui, pendant environ ..... jours ☐ Non

12. Est-ce que votre médecin a déjà diagnostiqué la borréliose chez vous dans le passé ?

.....

8.1.6 Questionnaire séances ultérieures

1. Quels sont les symptômes que vous présentez en ce moment? (Plusieurs réponses possibles)

☐ pas de symptômes **Symptômes générales:** ☐ Maux de tête ☐ Symptômes grippaux ☐ Raideur de la nuque ☐ Troubles d'équilibre ☐ Manque de concentration ☐ Troubles de la mémoire ☐ Fatigue exceptionnelle ☐ Insomnie ☐ Changement de la personnalité ☐ Dépression	**Neuroborréliose:** ☐ Névralgie ☐ Inflammation du cerveau ☐ Paralysie ☐ Troubles sensoriels **Arthrite:** ☐ Arthralgie Articulations affectées: ..... ☐ Douleurs musculaires	**Manifestations sur la peau :** ☐ Érythème migrant ☐ Décoloration rouge-bleuâtre ☐ Enflure ☐ Apparence craquelée rappelant un papier de cigarette froissé ☐ Peau anormalement mince ☐ Vaisseaux sanguins transparaissant ☐ Nœuds sur la peau **Autres symptômes:**

Si vous souffrez de manifestations de la peau, vous pouvez joindre une photo ou nous l'envoyer à l'adresse e-mail borreliose@crp-sante.lu **.**

2. Depuis quand observez-vous les symptômes susmentionnés?

.....

3. Thérapie

Médicament/Antibiotique: ..... Début de la thérapie: ..... Durée de la thérapie: .....

4. Commentaires supplémentaires

……....
.
……....
.
……
.
……
.

Veuillez répondre aussi aux questions suivantes, qui sont importantes pour une autre étude actuellement effectuée au Département d'Immunologie:

5. Avez-vous déjà eu la rougeole?

☐ Oui, il y a environ ..... an(s)

☐ Non

☐ Je ne sais pas

6. Êtes-vous vacciné contre la rougeole?

☐ Oui, j'ai été vacciné il y a environ ..... an(s)

☐ Non

☐ Je ne sais pas

8.2 Conference participations

11th - 12th September 2012	**Life Sciences PhD days** (Limperstberg, Luxembourg) Poster presentation: Exploring the Lymphocyte Repertoires of Lyme Borreliosis Patients as a Potential Tool for Epidemiology, Symptomatology and Diagnosis
28th November 2012	**SaarLorLux meeting** (Nancy, France) Oral presentation: Exploring the Lymphocyte Repertoires of Lyme Borreliosis Patients
15th - 16th November 2012	**Antigen processing and presentation in health and disease** (Dommeldange, Luxembourg)
25th - 27th February 2013	**AK B cell meeting** (Schluchsee, Germany)
13th - 14th June 2013	**AK Vakzine Meeting** (Freiburg, Germany) Oral presentation: Exploring the Lymphocyte Repertoires of Lyme Borreliosis Patients as a potential Tool for Epidemiology, Symptomatology and Diagnosis
18th - 21st August 2013	**13th International Conference on Lyme Borreliosis and Other Tick-Borne Diseases** (Boston, USA) Poster presentation: From prevalence studies to the development of novel diagnostic tests for Lyme Disease
02nd - 7th September 2013	**ESF-EMBO Symposium with support from EFIS – B Cells from Bedside to Bench and Back Again** (Pultusk, Poland) Poster presentation: Exploring the B cell Repertoires of Lyme Borreliosis Patients as a potential Tool for Epidemiology, Symptomatology and Diagnosis
09th - 10th September 2013	**Life Sciences PhD Days** (Limpertsberg, Luxembourg)

	Oral presentation: Exploring the B cell Repertoires in Lyme Borreliosis Patients towards Personalized Medicine
12th - 14th November 2013	**13th Euroconference on Clinical Cell Analysis** (Kirchberg, Luxembourg) Poster presentation: Exploring the B cell Repertoires in Lyme Borreliosis Patients towards Personalized Medicine
22nd - 28th March 2014	**Gordon Research Conference - Antibody Biology & Engineering" and "Antibody Biology & Engineering Gordon-Merck Research Seminar"** (Lucca, Italy) Poster presentation: Evaluation of Borrelia specific CDR3 sequences as potential biomarker for acute Lyme Borreliosis
11th September 2014	**17th Saar-Lor-Lux workshop** (Nancy, France) Oral presentation: Combining flow cytometry with next-generationsequencing for identification of antigen-specific B cells in Lyme Borreliosis patients
15th - 16th September 2014	**Life Sciences PhD Days** (Luxembourg, Luxembourg) Oral presentation: Combining flow cytometry with next-generationsequencing for identification of antigen-specific B cells in Lyme Borreliosis patients
17th - 20th September 2014	**DGfI meeting** (Bonn, Germany) Poster presentation: What can we read from the B cell repertoires of Lyme Borreliosis patients?
27th - 30th September 2015	**14th International Conference on Lyme Borreliosis and other tick borne diseases** (Vienna, Austria) Poster presentation: The promises and challenges of analyzing the B cell repertoire of acute Lyme disease patients
15th - 16th October 2015	**AK Vakzine Meeting** (Freiburg, Germany) Oral presentation: The promises and challenges of analyzing the B cell repertoire of acute Lyme disease patients
12th - 13th November 2015	**7th Annual Next Generation Sequencing Congress & 3rd Single Cell Analysis Congress** (London, United Kindom)
21st - 22nd September 2015	**Life Sciences PhD Days** (Belval, Luxembourg) Poster: Dissecting the B cell repertoire of acute Lyme disease patients
25th - 27th April 2016	**PhD School Luxembourg** (COST Action FA1207) (Esch-sur-Alzette, Luxembourg) Oral presentation: What I learned from my PhD besides scientific knowledge

8.3 Publications

Update of bookchapter: **Josiane Kirpach**, Claude P. Muller (2015), Epitopes. Encyclopedia of Life Sciences

Sally Cutler, Nataliia Rudenko, Marina Golovchenko, Wibke J. Cramaro, **Josiane Kirpach**, Sara Savic, Iva Christova, Ana Amaro (2016), Diagnosing Borreliosis. Vector Borne Zoonotic Dis.

Currently we have three manuscripts in preparation to which Josiane Kirpach has contributed and where she will be among the first few authors. Please find below the current running titles of the concerned manuscripts:

Detection of a low level and heterogeneous B cell immune response in peripheral blood of acute Lyme disease patients with next generation sequencing. **Kirpach** et al. in preparation.

Using *in vitro* stimulation to isolate memory B cell receptor sequences reactive towards common antigens and determination of their levels in B cell repertoire deep sequencing data. **Kirpach** et al. in peparation.

Next generation sequencing reveals a constrained viral quasispecies evolution under crossreactive antibody pressure. **Kirpach** et al. in preparation.

8.4 Acknowledgements

In the methods section I already mentioned who supported me in the different types of experiments, but I would like to take the opportunity here to thank all the people that contributed in some way to my PhD project.

First of all I would like to thank Prof. Muller for accepting me as PhD student in his laboratory and for supervising me during the last five to six years. This PhD allowed me to evolve in many different aspects. Thank you for that. A very big Thank you also to Anna Reye and Axel Dubois for their great support, especially at the beginning of my project.

During the last few years, Alessia Colone was a geat MarLyBor team-mate. She did not only provide moral and strategic support, but she taught me also many life lessons and a lot about labwork. She supported me in all kinds of experiments including single cell PCRs and ELISAs. She performed the *in vitro* single cell stimulation experiments, which were very valuable for this project. It was really great to have you on board for the last few years.

A big thank you also to Wibke Jochum. It was really great to have you as MarLyBor team-mate as well.

Lynn Wenandy also supported me during a transient period of my project. She helped with the selection of the markers for our multicolor B cell panel.

René Brons was a great support for the setup and execution of FACS experiments. Thank you very much for fruitful discussions and for helping to troubleshoot whenever I had problems with a FACS protocol. A big thank you also for teaching me about multicolor flow cytometry and the Kaluza data analysis software. It was always a great pleasure to have discussions with you.

During my PhD I learned how expensive research is, so that I also really appreciate the nice financial support we got from the Luxembourg National Research Fund (AFR PhD grant N°3986187) and from the Luxembourgish Ministry of Research (project Marlybor, REC-LNSI-20110308). This allowed me to use cutting-edge technologies and to go to many exciting meetings.

I would like to thank all the patients that sacrificed their precious time and blood samples for this study. Also many thanks to the doctors who took the time and initiative to send their patients to us.

I would want to thank Laboratoires Réunis, our Clinical and Epidemiological Investigation Center and the Inspection Sanitaire for their great support in sample collection.

Although you can probably not remember me, I would like to thank Claudia Giesecke and Christian Busse with whom I had my first discussions about B cells on my first meeting. Your feedback helped me to know in which direction I would want to go with this project. These were not the only people with whom I had exciting discussions at meetings and I would like to thank all of those who took their time to answer my questions and who showed their interest in me eventhough I am(was?) only a little PhD student.

As for the setup and execution of protocols I also got great support from many different people that I would like to mention here. Thank you to Patil Sarita, Roland Buelow and Michael Osborn for sharing their experiences with single cell protocols. Then a big thank you to Elisabetta Traggiai and John Lindner (Department of Mechanistic Immunology, Novartis Institutes for BioMedical Research, Basel, Switzerland) for sharing their *in vitro* stimulation protocol with our laboratory. I also have to thank Jean-Philippe Buerckert for putting my attention to the nice paper by Vollmers et al and for sharing his initial library preparation protocol.

I also got great support for the data analysis part. Jacob Galson and Julia Bischof were so kind to share their sequence clustering scripts with me. Deborah Dunn-Walters and David Kipling explained me how to determine spectratype distortion. Then I also have to thank Anke Wienecke-Baldacchino who wrote the initial deep sequencing pre-processing pipeline with the support from Axel Dubois and Jean-Philippe Buerckert. Jean-Philippe Buerckert and William Faison took over these tasks later and performed the data pre-processing part for this study. They also implemented the Change-o pipeline from Steven Kleinstein's group in our laboratory. William Faison took over the clustering analysis with this tool. Jean-Philippe introduced me to bash scripting.

A very huge thank you to I-Na for her moral and strategic support. I am really greatful for that. I don't know what I would have done without you.

Thank you also to Anne Molitor, Sandra Gohrbandt and Regina Sinner for setting up the Ion Torrent technology in our laboratory. With this regard I would like to specially thank Regina Sinner for teaching me on how to use the Ion Torrent machine. At the same time I would like to thank Nastasja Hauck for her nice support during our first Ion Torrent runs that we had to do alone without Regina. I am really happy to look back at the many fruitfull and motivating discussions we had, Nastasja.

A big thank you also to Sophie Farinelle, who helped me with the setup of the peptide ELISA and who allowed me to use her ELISA buffers.

Then I would also want to thank Vitor Lopez and Mark Leclerc for their always shiny and happy attitudes and helpfulness. It is really nice to have so friendly people in the institute.

I would also want to thank the technicians of our laboratory who were always helpful when I had questions and who made sure that our laboratories are running smoothly: Aurélie Sausy, Sophie Mériaux, Claire Dording, Stephanie Wolter, Stéphanie Willieme, Sophie Farinelle, Stéphanie Schmitz, Emilie Charpentier and Martyna Marynowska.

Even if the cloning and expression experiments were not really successful and we did not have the time to finish this part of the project, I would like to thank Kenza Benkirane and Anouk Even for their great support on this part. You both have been amazing students.

A big thank you also for the nice office and lunch discussions with Anna, Axel, Alessia, I-Na, Wibke, Nastasja, Ni, Martha, Silvia, Sonja, Maude, Lisa and all the others that occasionally joined us for lunch. I met really nice people in this institute.

Then I would also want to thank the IT team from the Luxembourg Institute of Health as well as Oliver Hunewald. Thank you for your patience and help whenever I had questions or problems with my computer or data analysis.

Carole Weis and Ulla Muller were a great support for all administrative and organizational parts.

Then I would like to thank Chris Kaufmann (and his family), who was a nice partner during the first years of my PhD. A big thank you also to my parents, my two brothers and my closest friends. Thanks for your patience and for listening to my many problems. I am really grateful for your moral support.

I hope I did not forget someone. My appologies if I did so! I wish all of you only the best for your future!