The role of the Epstein Barr Virus in B cell tolerance loss in rheumatoid arthritis disease

Edoardo Prediletto

Submitted in partial fulfilment of the requirements of the Degree of Doctor of Philosophy

Queen Mary University, Barts and the London School of Medicine and Dentistry, William Harvey Research Institute

2020

ACKNOWLEDGMENTS

Firstly, I would like to express my sincere gratitude to my advisor Prof. Michele Bombardieri for the continuous support of my PhD study and related research, for his patience, motivation, and immense knowledge. His guidance helped me in all the time of research and writing of this thesis. Besides my advisor, I would like to thank Dr. Thorley Lawson for his insightful comments and encouragement, who had actually turned the tide during the project. I would like also to thank all my colleagues, my family and my partner for their continuous support.

STATEMENT OF ORIGINALITY

I, Edoardo Prediletto, confirm that the research included within this thesis is my own work or that where it has been carried out in collaboration with, or supported by others, that this is duly acknowledged below, and my contribution indicated. Previously published material is also acknowledged below.

Dr Paola Migliorini and Dr Federico Pratesi performed the ELISA for VCP1 and 2, HCP1 and 2. Dr. Elena Pontarini thought me and performed part of the flow cytometry analysis for the optimization of the co-culture system.

Part of the 10X single cells sequencing experiments has been performed at the Genome Center (QMUL) as part of a service, specifically the loading of samples on the Chromium system and the Illumina Sequencing.

Dr. Thorley Lawson actively contributed to the project development.

Dr. Michele Bombardieri had overall supervision of the work.

I attest that I have exercised reasonable care to ensure that the work is original and does not to the best of my knowledge break any UK law, infringe any third party's copyright or other Intellectual Property Right, or contain any confidential material.

I accept that the College has the right to use plagiarism detection software to check the electronic version of the thesis.

I confirm that this thesis has not been previously submitted for the award of a degree by this or any other university.

Signature:
Date: **01/02/2021**

Details of collaboration and publications:

This work has been presented at:

British society of immunology conference, Liverpool, UK, 2016

British Society of immunology conference, Brighton, UK 2017

International union of immunological society conference, Beijing, China, 2019

British society of immunology conference, Liverpool, UK, 2019

This work was supported by the Versus Arthritis: Grant code 21268

ABSTRACT

Although many efforts have been done in understanding the aetiology of rheumatoid arthritis (RA), the Epstein-Barr Virus (EBV) remains one of the strongest candidates to be investigated as possible environmental factors in inducing RA. Of relevance, a high percentage of RA patients are characterised by a strong immune response to EBV, with high blood DNA-viral load and circulating cross-reactive (auto)antibodies reacting against self-citrullinated antigens (ACPA) and viral proteins.

Due to its lymphotropic behaviour, EBV might be able to rescue an autoreactive B cell phenotype and thus eludes the immune system checkpoints. Also, the virus and the infected B cells have been proposed as a possible source of citrullinated epitopes triggering the disease.

To obtain long-standing EBV-infected B cells, I setup an in vitro co-culture system whereby RA fibroblasts-like synoviocytes (RA-FLS) obtained from synovial fluid and synovial tissue of patient were cultured with CD19+B cells for 28 days. RA-FLS were used as feeding environment for the CD19+B cells obtained from ACPA+ RA patients and healthy donor. CpG was added to induce plasmacells differentiation to test the presence of ACPA antibodies. B cells were then recovered, and analysed trough fluorescence activated cell sorting (Facs). Finally, molecular biology analysis has been performed to detect and quantify specific EBV-related gene such as the BamH1-W repeat region. Also, RA-FLS were recovered in order to investigate any possible modification induced by the CD19+B cells in the system.

Preliminary data suggested that the EBV+ CD19+B cells population has a higher proliferation rate when incubated with RA-FLS, although this is not exclusive of RA patient since such proliferation was detected also in healthy donor. Nevertheless, this might reflect an in vivo mechanism in

which the EBV+ CD19+B cells from the peripheral compartment might find a preferential proliferating niche in the synovium, due to the RA-FLS.

Furthermore, I am proposing a new method to obtain naturally EBV+ RA-CD19+B cells in order to develop a better tool to study the role of the virus in RA pathogenesis.

Table of Contents

LIST OF FIGURES

LIST OF TABLES

ABBREVIATION

A	Alanine
ACIF	Anti-complement immunofluorescence
ACPA	Anti- citrullinated protein antibodies
ACR	American college of rheumatology
AKA	Anti-keratin antibodies
APC	Antigen-presenting cells
APF	Anti-perinuclear factors
BCL	B-cell lymphoma protein
BL	Burkitt lymphoma
CCL2	Chemokine (C-C motif) ligand 2
CCL5	Chemokine (C-C motif) ligand 5
CCR7	Chemokine (C-C motif) receptor 7
CNS	Central nervous system
CRT	Calreticulin
CSR	Class-switch recombination
CTLA 4	cytotoxic T-lymphocyte-associated protein 4
CXCL10	C-X-C motif chemokine ligand 10
CXCL12	C-X-C motif chemokine ligand 12
CXCL13	C-X-C motif chemokine ligand 13
CXCR4	C-X-C motif chemokine receptor 4
CXCR5	C-X-C motif chemokine receptor 5
D	Aspartic acid
DAF	Decay-accelerating factor
DAMP	Damage associated molecular patterns
DAS28	Disease activity joints 28
DC	Dendritic cells
DN	Double negative
DNMT1	DNA methyltransferase 1
dsDNA	Double strand DNA
dsRNA	Double strand RNA
DX	Diagnosis
E	Glutamic acid
EA	Early antigen
EBER	Epstein–Barr virus-encoded RNA
EBNA	Epstein–Barr nuclear antigen

EBNA-LP	Epstein-Barr nuclear antigen leader protein
EBV	Epstein Barr virus
EGC	Ectopic germinal center
ELS	Ectopic lymphoid structure
EMR	Experimental Medicine and Rheumatology
EMR-RTB	EMR Research Tissue bank
ENK	Extranodal natural killer
EULAR	European leagues against rheumatism
FACS	Fluorescence-activated cell sorting
FAP	Fibroblast activation protein
FC	Fragment crystallizable region
FDC	Follicular dendritic cells
FISH	Fluorescence *in situ* hybridization
FITC	Fluorescein isothiocyanate
FS	Felty's syndrome
GC	Germinal center
GWAS	Genome wide association studies
GYPA	Glycophorin A
H	Heavy chain
HCP1	Histone citrullinated peptide 1
HCV	Hepatitis C virus
HEV	High endothelial venules
HierBAPS	Hierarchical bayesian analysis of population structure
HIV	Human immunodeficiency virus
HL	Hodgkin lymphoma
HLA	Human leukocyte antigen
HMGB1	High mobility group Box 1
HRS	Hodgkin–Reed–Sternberg
Ig	Immunoglobulin
IL-12p70	Interleukin 12 p70
IL-2	Interleukin-2
IL-6	Interleukin-6
IM	Infectious mononucleosis
ISH	In situ hybridization
K	Lysine
L	Light chain
LCL	Lymphoblastoid cell lines
LFA-3	Lymphocyte function-associated antigen 3
LMP	Latent membrane protein
LPS	Lipopolysaccharide
LT-α	Lymphotoxin-alpha

LT-β	Lymphotoxin-beta
LT-β	Lymphotoxin-beta receptor
MADCAM1	Mucosal vascular addressin cell adhesion molecule 1
MHC	Major histocompatibility complex
MS	Multiple sclerosis
N/A	Data not available
NET	Neutrophil extracellular traps
NIMA	Non inherited maternal antigen
NLR	Nod-like receptors
NPC	Nasopharyngeal carcinoma
PAMP	Pathogen-associated molecular patterns
PCA	Principal component analysis
PKR	Protein Kinase RNA-activated
PTLD	Post-transplant lymphoproliferative disease
Q	Glutamine
QC	Quality control
QMUL	Queen Mary University of London
Qmul-Rbb	Qmul- Research biobank
Q-PCR	Quantitative PCR
R	Arginine
RA	Rheumatoid arthritis
RA-FLS	Rheumatoid arthritis fibroblast like synoviocytes
RANA	Rheumatoid-arthritis-associated nuclear antigen
RF	Rheumatoid factor
RIG-I	Retinoic acid-inducible gene I
SCID	Severe combined immunodeficient mice
SF	Synovial fluid
SHM	Somatic hypermutation
SLE	Systemic lupus erythematosus
SLO	Secondary lymphoid organs
SS	Sjögren syndrome
SSA/Ro	Sjögren's syndrome antigen A (ribonucleoprotein autoantigen)
SSB/LA	Sjögren's syndrome antigen B (La autoantigen)
STRING	Search tool for the retrieval of Interacting genes/proteins
T1DM	Diabetes mellitus type 1
TCDD	2,3,7,8-tetrachlorodibenzo-p-dioxin
TLR	Toll-like receptors
TLO	Tertiary lymphoid organs
TNFR	Tumour necrosis factor receptor
TRADD	TNF receptor-1 associated death domain
TRAFF	TNF receptor associated factors

Treg	T regulatory cells
TSNE	t-distributed stochastic neighbour embedding
UMAP	Uniform Manifold Approximation and Projection
UMI	Unique molecular identifiers
VCA	Viral capsid antigen
VCP1	Viral citrullinated peptide 1
VH	Variable heavy chain
VL	Variable light chain
XLP1	X-linked lymphoproliferative disease type 1

CHAPTER 1 INTRODUCTION

A complex autoimmune disease: Rheumatoid Arthritis

Overview

Rheumatoid arthritis (RA) is the most common inflammatory erosive polyarthritis affecting 1% of the general population and half-million of people in the UK. The disease is characterized by a chronic inflammation which affects mainly the synovial joints and leads to strong synovial alterations (Fig.1.1).

At the peripheral blood level one of the hallmarks of the disease is an autoreactive B cell dysregulation that can results in the production of autoantibodies, such as those against citrullinated proteins (ACPA) which are present in around 70% of RA patients.

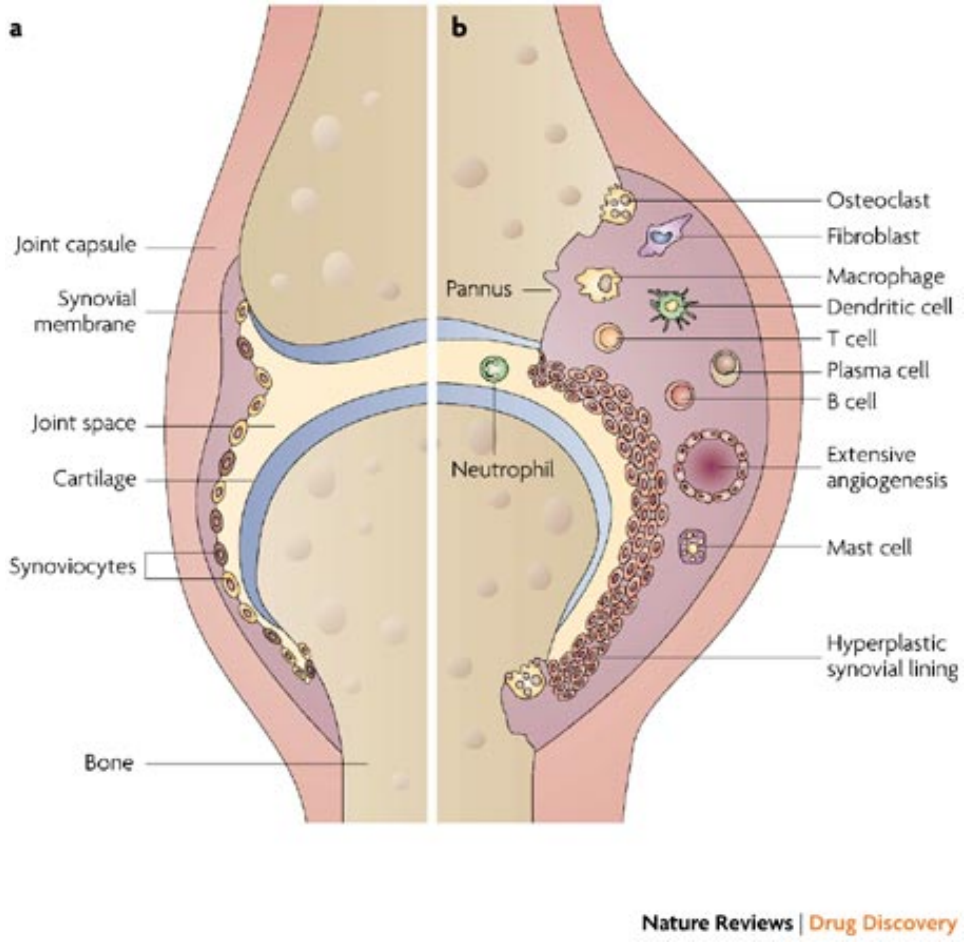


Figure 1.1 Normal joint compared to rheumatoid arthritis joint. In a healthy joint (a) the synovial membrane appears thin with a single cell layer. In a rheumatoid one (b) cells in the synovial membrane become activated and hyperproliferative, increasing the thickness and developing into the characteristic "pannus", which then interfere with bone and cartilage. From: Strand, V., R. Kimberly, and J.D. Isaacs, *Biologic therapies in rheumatology: lessons learned, future directions.* Nat Rev Drug Discov, 2007. **6**(1): p. 75-92. [1].

Classification criteria

Referring to Dr Wasserman's paper, in 2010 the American College of rheumatology (ACR) and the European Leagues against Rheumatism (EULAR) created new classification criteria for RA, aiming to diagnose the disease earlier in patients who may not meet the older 1987 ACR classification. The most important changes have been the exclusion of features present in well-established RA but not in early RA cases, such as the presence of nodules or radiographic erosive changes, together with the evidence of symmetric arthritis [2]. A comprehensive list of these updated criteria is available in Table 1.1. It is important to mention the contribution from Dutch researchers whom analysis system has been developed in order to predict and identify those patients with undifferentiated arthritis who will, eventually, develop RA (Table 1.2).

Clinical aspects

Pain and stiffness in multiple joints are the typical clinical presentations in RA patients. Most involved joints are the wrists, the proximal interphalangeal and metacarpophalangeal of the hands. Also, the metatarsophalangeal joints of the feet are commonly targeted. Through the DAS28 – where DAS stands for "disease activity" and 28 refers to the joints that are examined in this assessment - the disease activity can be monitored from baseline through the radiological progression [3]. Visual and tactile joint examination may reveal swelling and thickening of the synovia due to synovitis. Patients usually present also with prolonged morning stiffness, weight loss and systemic symptoms of fatigue. For a detailed list of extra-articular symptoms refer to Table 1.3.

	Score
Target population (who should be tested?): patients who	
1) have at least one joint with definite clinical synovitis (swelling)*	
2) with the synovitis not better explained by another disease†	
Classification criteria for RA (score-based algorithm: add score of categories A through D; a score of ≥ 6 out of 10 is needed for classification of a patient as having definite RA)‡	
A. Joint involvement§	
One large joint‖	0
Two to 10 large joints	1
One to three small joints (with or without involvement of large joints)¶	2
Four to 10 small joints (with or without involvement of large joints)	3
> 10 joints (at least one small joint)**	5
B. Serology (at least one test result is needed for classification)††	
Negative RF *and* negative ACPA	0
Low positive RF *or* low positive ACPA	2
High positive RF *or* high positive ACPA	3
C. Acute phase reactants (at least one test result is needed for classification)‡‡	
Normal CRP and normal ESR	0
Abnormal CRP or normal ESR	1
D. Duration of symptoms§§	
< six weeks	0
≥ six weeks	1

ACPA = anti-citrullinated protein antibody; CRP = C-reactive protein; ESR = erythrocyte sedimentation rate; RA = rheumatoid arthritis.

**—The criteria are aimed at classification of newly presenting patients. In addition, patients with erosive disease typical of RA with a history compatible with prior fulfillment of the 2010 criteria should be classified as having RA. Patients with long-standing disease, including those whose disease is inactive (with or without treatment), who, based on retrospectively available data, have previously fulfilled the 2010 criteria should be classified as having RA.*

†—Differential diagnoses differ in patients with different presentations, but may include conditions such as systemic lupus erythematosus, psoriatic arthritis, and gout. If it is unclear about the relevant differential diagnoses to consider, an expert rheumatologist should be consulted.

‡—Although patients with a score of less than 6 out of 10 are not classifiable as having RA, their status can be reassessed, and the criteria might be fulfilled cumulatively over time.

§—Joint involvement refers to any swollen or tender joint on examination, which may be confirmed by imaging evidence of synovitis. Distal interphalangeal joints, first carpometacarpal joints, and first metatarsophalangeal joints are excluded from assessment. Categories of joint distribution are classified according to the location and number of involved joints, with placement into the highest category possible based on the pattern of joint involvement.

‖—"Large joints" refers to shoulders, elbows, hips, knees, and ankles.

¶—"Small joints" refers to the metacarpophalangeal joints, proximal interphalangeal joints, second to fifth metatarsophalangeal joints, thumb interphalangeal joints, and wrists.

***—In this category, at least one of the involved joints must be a small joint; the other joints can include any combination of large and additional small joints, as well as other joints not specifically listed elsewhere (e.g., temporomandibular, acromioclavicular, sternoclavicular).*

††—Negative refers to international unit values that are less than or equal to the upper limit of normal for the laboratory and assay; low positive refers to international unit values that are higher than the upper limit of normal but three or less times the upper limit of normal for the laboratory and assay; high positive refers to international unit values that are more than three times the upper limit of normal for the laboratory and assay. When rheumatoid factor information is only available as positive or negative, a positive result should be scored as low positive for rheumatoid factor.

‡‡—Normal/abnormal is determined by local laboratory standards.

§§—Duration of symptoms refers to patient self-report of the duration of signs or symptoms of synovitis (e.g., pain, swelling, tenderness) of joints that are clinically involved at the time of assessment, regardless of treatment status.

Table 1.1. The 2010 American College of Rheumatology/European League Against Rheumatism Classification Criteria for RA. From *Diagnosis and Management of Rheumatoid Arthritis*, *Amy Wasserman, 2011*, readapted from *2010 rheumatoid arthritis classification criteria: an American College of Rheumatology/European League Against Rheumatism collaborative initiative, Aletaha D, Neogi T, Silman AJ, et al. Ann Rheum Dis. 2010* [4].

Patient characteristics	*Points*
Age	Years × 0.02
Female sex	1.0
Distribution of affected joints (patients may receive points for more than one item)	
Small joints of hands or feet	0.5
Symmetric	0.5
Upper extremities	1.0
Upper and lower extremities	1.5
Score for morning stiffness on a 100-mm visual analog scale	
26 to 90 mm	1.0
> 90 mm	2.0
Number of tender joints	
Four to 10	0.5
≥ 11	1.0
Number of swollen joints	
Four to 10	0.5
≥ 11	1.0
C-reactive protein level	
5 to 50 mg per L (47.62 to 476.20 nmol per L)	0.5
≥ 51 mg per L (485.72 nmol per L)	1.5
Positive for rheumatoid factor	1.0
Positive for anti-citrullinated protein antibody	2.0
Total:	_____

*Score**	*Number with RA*	*Number without RA*	*Likelihood ratio*	*Percentage with RA at one year*
0 to 3.5	0	109	0	0
3.51 to 6.5	41	214	0.42	16
6.51 to 8.5	71	53	3.0	57
≥ 8.51	63	11	12.7	85

RA = rheumatoid arthritis.

**—Score rounded to the nearest number ending in 0 or 0.5.*

Table 1.2. Clinical rule for predicting the Risk of RA Patients with undifferentiated arthritis. From *Diagnosis and Management of Rheumatoid Arthritis, Amy Wasserman, 2011*, readapted from *A prediction rule for disease outcome in patients with recent-onset undifferentiated arthritis, Arthritis Rheum, van der Helmvan Mil, 2007,* with scoring information from reference 5 [2] [5, 6].

ORGAN	Clinical presentation
Skin	Vasculitis, nodules
Cardiac	Pericarditis, myocardial fibrosis/granulomatous disease
Pulmonary	Interstitial pneumonitis, rheumatoid nodules, serositis
Neurological	Mononeuritis multiplex, peripheral neuropathy
Haematological	Anaemia, Felty's syndrome
Exocrine	Secondary Sjögren's syndrome
Others	Amyloidosis, systemic vasculitis, osteoporosis

Table 1.3. Extra-articular manifestations in rheumatoid arthritis.

Epidemiology

Based on the most recent analysis, RA is estimated to affect between 0.24 and 1% of the global population. More specifically the incidence of the disease in Europe and United States (where it has been more thoroughly studied) is estimated to be approximately 40 per 100'000 persons with women having usually twice the chance of getting RA compared to men (lifetime risk: women 3.6% vs men 1.7) [7] [8]. Of notice the Pima Native Americans population has a 10.3 times higher risk of developing the disease than the global population [9]. In Nigeria and in less developed parts of Africa, RA is almost undetectable, whereas prevalence rates rise to nearly 1% in black populations in urban areas, further suggesting some sort of environmental factor in the etiology of the disease.

Etiology

Similar to many autoimmune diseases, RA has several levels of complexity with a multifactorial etiology. There is evidence that both genetic and environmental factors are involved.

Genetic Factors and the Shared Epitope hypothesis in RA

RA has been widely studied in familial clustering and, specifically, in twin studies (monozygotic VS dizygotic). These large genetic screening in North Europe have allowed to find a concordance rates for RA of 12.3% and 15.4% in monozygotic twins, while these values dropped to 3.5% in dizygotic twins. This has led to a final estimation of heritability for RA around 65% [10].

In order to understand and identify possible susceptibility loci in 1989 Deighton et al have been performed for RA whole genome scans using multi-locus nonparametric linkage analysis. The final result showed that

human leukocyte antigen (HLA) region was the only one with statistical significance in term of linkage. Other authors have then connected the HLA to an higher risk, accounting for 30%-50% of all genetic susceptibility in RA [10, 11]. This was further refined by Stastny, who hinted a stronger association with a specific HLA gene, the β-chain of the HLA-DR molecule, expressed on the surface of antigen-presenting cells (APC). In the late 1990s Gregersen et al managed to completely sequence the HLA-DRB1 loci, and, for the first time, they proposed the "shared epitope (SE) theory", trying to explain the association between RA and the HLA-DR: several alleles present in the HLA-DRB1 would account for susceptibility alleles trough a homologous amino acid sequence in the third hypervariable region of the first domain of the molecule [12]. At a molecular level, the region of interests spans from position 70 to position 74, having either a glutamine (Q), a Lysine (K) or an arginine (R) at 70 and Alanine(A) at 74 ($_{70}$QRRAA$_{74}$ or $_{70}$KRRAA$_{74}$ or $_{70}$RRRAA$_{74}$). Specifically, these aminoacid sequence can be encoded by several DR: the HLA-DR4 (DRB1*0401, *0404, and *0405), HLA-DR1 (DRB1*0101 and *0102), and HLA-DR10 (DRB1*1010) and this motif constitute an α-helical domain creating one side of the antigen binding pocket. This has led researchers to speculate that the mechanism of association might rely on 1) antigen presentation of arthritogenic antigens [13], 2) the selection of T cell repertoire [14] 3) a disrupted generation of functional T regulatory cells (Treg) [15]. In contrast Zanelli et al described that sequences having aspartic acid (D) and glutamic acid (E) in position 70 and 71 ($_{70}$DERAA$_{74}$) which is expressed by certain HLA-DR alleles (DRB1*0103, *0402, *1102, *1103, *1301, *1302) have actually a protective effect. They also refined the haplotypes connected to RA by showing that susceptibility to the disease is not exclusively related to DR alleles alone but more likely to both HLA-DQ-DR alleles [16]. A more thorough classification of the HLA-DR SE component has been proposed by Tezenas du Montcel in 2005, where she divided the alleles in 3 subgroups (S2, S3P, and L) according to their susceptible/protective phenotype [17].

It has been to be mentioned that the SE association it is not exclusive of RA but has been found in other autoimmune diseases, such as type I diabetes (T1DM) [18]. Adding more complexity, around 20% of RA patients are SE negative, making a direct connection less plausible. However a study have shown a possible link between this negativity and HLA-DR4 coding non-inherited maternal antigens (NIMAs) in women with the onset of the disease in child-bearing age (<45 years), but not in older women or in men [19]. The possible mechanism could be explained by maternal and fetal microchimerism, a phenomenon that takes place during pregnancy and in which the mother and the foetus can interchange cells and soluble antigens, developing tolerance. Intriguingly maternal microchimerism has been associated also to autoimmunity [20].

It is fascinating that the presence of these alleles has been strongly related to the production of autoreactive antibodies, which are present in 70% of RA patient. Indeed, ACPA seropositivity is more strongly related to SE-positive patient and rarely detected in SE-negative [21], with allele-dose effect on antibody production: Snir et al. showed that RA patients carrying double SE alleles produce higher levels of ACPA (specifically anti-CCP antibodies, anti-citrullinated fibrinogen and anti-CEP-1) compared to those patient who are carrying a single SE or no allele at all [22, 23].

It has since then been hypothesized that this strong association might be the evidence of a SE-restricted presentation of citrullinated of antigens model. Nevertheless, despite these results, a direct connection between the SE and the specific peptide binding and presentation of citrullinated antigens or restricted T cell response to citrulline has been difficult. These topics are further expanded in the “Autoantibodies in RA: ACPA and RF” and in “EBV molecular mimicry” chapters.

All these connections have led to the speculation of several etiological model for RA based on this peculiar mechanism. An example can be the fact that the QKRAA sequence expressed by the HLA-DRB1*0401 shares sequence homology with EBVgp110, which is usually targeted by the immune response during infection [24].

It has to be mentioned that other genes have been found to be associated with RA: Genome wide association studies (GWAS) and single nucleotide polymorphism analysis (SNP) have highlighted elements such as ligands and transcriptional factors (CD40 or STAT4) [25], molecules involved in interaction between T and APC such as cytotoxic T-lymphocyte-associated protein 4, (CTLA4), chemokines (CCL21) or signals related to the activation of T cells (PTPN22) [26]**.** Nevertheless, they represent less than 5% of the overall genetic contribution compared to HLA genes.

Environmental factors in autoimmunity and in autoinflammatory diseases: a wider perspective for RA:

Current evidences and studies have shown that genetic predisposition can account for only a 30% of autoimmune diseases. Recent works have shown that the environment plays a much more fundamental role in triggering and shaping an autoimmune response at a given specific period of time [27]. It is very likely that the presence of both genetic and multiple environmental factors it is necessary for the break of immune tolerance which then result in the immune system attacking the host. Most of these factors have been elegantly described by Anaya et al in the review "The Autoimmune Ecology", and some of them will be reported here. This work defines the term "exposome", which is the whole range of factors to whom a subject can be exposed during the lifespan that can lead to autoimmunity. These can be both internal (e.g. viruses or bacteria) and external (e.g. lifestyle, climate, exposure to contaminant, air pollution) and interact at different levels of the immune system, such as the innate or adaptive immunity, or trough epigenetic and post-translational modifications. For some other factors we still do not have functional mechanisms but strong epidemiological association. Some of the most important will be summarised here but a complete list is beyond the purpose of this thesis and can be consulted on the review [28].

Toll like receptors (TLRs), as essential crossroads for the innate system, are important in dealing with treats, defined as PAMPs (pathogen-associated molecular patterns) or DAMPS (damage-associated molecular patterns). Upon activation, TLRs can upregulate chemokines and cytokines and define the duration and the extent of the inflammatory response. For their role, overactivation of these pathways has been often hypothesized to be linked to autoimmunity [29]. High mobility group box 1 (HMGB1) - a protein which is usually localized inside the cells - can act as a DAMP when released and can work in synergy with PAMPs such lipopolysaccharide (LPS), double stranded DNA or RNA (dsDNA, dsRNA), eventually inducing expression of prostaglandin that promote inflammation in arthritis [30]. Another example are peptidoglycans, which can be released from bacteria and then bind nod like receptors (NLRs) or TLR2. Particularly, it has been found to be increased in cells isolated from RA synovial tissue, such as macrophages or DCs [31]. Molecule released from dying cells upon ultraviolet irradiation have been reported to induce DAMP and being connected to immune responses. Also, in SLE, nuclear material and dsDNA released from necrotic cells can also become self-antigens, and activate the TLR9 signal cascade which leads to autoantibodies production [32]. In addition, organ culture from synovial biopsies of RA patients have been found to express TLR2 and TLR 4 and to release TLR endogenous ligands in a process that can contribute to inflammation and arthritis [28].

Regarding the adaptive system, failure in the numerous checkpoints of B cell activation can leads to possible autoimmune mechanisms. Errors during the somatic hypermutation (SHM) can break self-tolerance leading to the production of autoreactive clones which can – if not inactivated - either produce autoantibodies or present autoantigens to T cells and produce inflammatory cytokines [33]. Estrogen and prolactin, two important sex hormones, have been reported to directly activate autoreactive B cells. Indeed, sex hormones are investigated as one of the main culprits in SLE for the predominance observed in female [34].

Furthermore, it has also been shown that women that undergoes hormonal replacement therapy have a higher odd of developing the disease [35]. As SLE, Sjogren Syndrome (SS) is also highly prevalent in women: males affected by the disease have usually less anti-ro antibodies and Raynaud's phenomena compared to the other sex [36], but they have an higher risk of lymphoma and neurological disorder [37].

Another element that can activate B cells is 2,3,7,8-tetrachlorodibenzo-p-dioxin (TCDD), an environmental factor released by incineration of several material - one these to be cigarettes smoking - and that has been widely associated with autoimmunity: animal model exposed to TCDD during prenatal development showed an increase in autoantibodies and autoimmune kidney lesions [38].

Regarding T cells, the most studied case in autoimmunity regard a subset of T helper cells that strongly participate in autoimmune conditions, which are the Th17. There are strong connection between the development of this subset and interaction with commensal bacteria in the intestine, all leading to a strong immune response and autoimmunity [39].

At the epigenetic level the best studied case in autoimmunity is probably SLE, where several environmental agents - cigarette smoke, mercury ultraviolet light, viral infections and medication - are known to induce oxidative stress, resulting in a reduction of DNA methyltransferase 1 (DNMT1) and consequent DNA methylation in CD4+ T cells that promote autoimmunity [40].

Several other environmental factors have been connected to this topic based on a wider epidemiological association. As previously mentioned, cigarette smoking is of those. It has been associated with an higher odd of developing SLE [41] and has been connected to seropositive RA (ACPA/Rheumatoid factor) in Caucasian and Latina Americans [42]. Smoking can also be directly linked to the SE and specific gene-environment interactions [43, 44] . For instance, it has been demonstrated that smoking can induce protein citrullination in lung macrophages among persons with the SE [45]. This hypothesis is reinforced by a study in which

– using HLADRB1*0401 transgenic mice – it has been reported that citrullination of certain peptides rather than other have made them more prone to be mounted and presented on HLA class II molecules with the HLA-SE, leading to a very specific immune response to citrullinated self-antigens [46] .

Also, habits such as alcohol and coffee consumption have showed association with seropositive RA [47].

Vitamin D deficiency - a crucial molecule for the immune system - has been associated with several autoimmune diseases, such as multiple sclerosis (MS), SLE, RA, inflammatory bowel disease and T1DM. These disease become more common as the distance from the equator increases and this has been related to diminished sun exposure [28].

Regarding organic solvents, chronic exposure to these potentially harmful agents has been associated as well with MS [48] and systemic sclerosis [49].

The economic status might be a factor of influence as well in this topic. It has been reported that in the US, African American develop MS at a younger age compared to European and Hispanic. This was matching with lower socioeconomical position that can result in difficulties in accessing neurology consultation [50].

The last environmental factors to be covered are microbial agents. The concept of a microbe triggering an autoimmune disease has been proposed since the 1870s [51], leading to an enormous amount of research in this direction, very often characterized by conflictual results. In particular, in RA, it does not look like a single microorganism can trigger the disease but rather a combination of bacterial and viral factor [52]. Also, it might be that a certain microorganism can trigger the disease only in a subset of cases or in cooccurrence with other clinical manifestation, such as psychological stress or microtrauma [53]. A crucial conceptual difference - as defined by Arleevskaya – is between the acquired and the innate hypothesis. In the first case the single event of infection might lead to autoimmunity, while in the second one, a genetical predisposition would make a certain group of

patients more susceptible to a specific microorganism [54]. Among the group relate to the acquired hypothesis the most interesting are Porphyromonas gingivalis (*P. gingivalis*) [55] which, as discussed later, is responsible for periodontal disease and can directly citrullinate human proteins.

Regarding those included in the innate hypothesis there are: Proteus mirabilis (*P. mirabilis*) [56], Mycobacterium tuberculosis [57], Mycoplasma [58], human retrovirus 5 [59], alpha virus [60], Cytomegalovirus [61], Rubella virus [62], parvovirus B19 [63], and the ubiquitous Epstein Barr Virus (EBV) [64]. The latter has been extensively and yet inconclusively investigated since the 1980's as a contributor to RA and is the focus of my PhD project. The virus has a strong connection not only to RA but to autoimmunity in general and specifically in MS and SLE. This topic will be discussed on the EBV chapter.

So far has been difficult to connect a single infectious agent as the cause of RA due to lack of consistent data and large microbial screening. Despite these controversies, new evidences are emerging regarding a novel role for EBV in the disease as I will discuss in the relevant section.

Immunopathogenesis in RA

As above mentioned, one of the hallmarks of the disease is a strong alteration of the synovial architecture. Stromal cells, in particular fibroblast like synoviocytes (FLS), can be activated and they are usually responsible for tissue invasion and matrix degradation. Contextually, the infiltration in the synovia of different immune cells subsets, such as B cells, T cells and macrophages can be found. Histologically, the difference in the grade of infiltration has been developed to characterize and stratify the disease in further subsets. Indeed several related works have been published by Pitzalis, Humby, Riberio and Lewis identifying three different" pathotypes" [65-67]: 1) lympho-myeloid, with abundance of B cells and plasmacells 2) diffuse -myeloid, with a predominance of myeloid cells and

a low percentage of B cells 3) pauci-immune, characterised by a prevalence of stromal cells and virtually no immune cells infiltrate.

Among those patients with clear synovial immune cell infiltration: i) 50% of the cases are characterized by predominant monocytes and macrophage but also T cells; ii) 25% expose sign of B cells and T cells without a clear organization; iii) the last 25% present with organized tertiary lymphoid organs (TLO), also called ectopic lymphoid structure (ELS), where B and T cell follicles are further clustered thanks to presence of CD21+ follicular dendritic cells (FDCs) [68]. Interestingly, this peculiar organisation has been described also in other autoimmune or inflammatory conditions as well as different type of solid cancers: the most important have been listed in Table 1.4 (the complete list can be consulted on the review from 2014 Pitzalis et al [69]).

Condition	Location	Antigen recognized by cells in ELS/TLO	Other known function of ELS/TLO in disease
Autoimmune Diseases			
Rheumatoid arthritis	Synovial tissue	Rheumatoid factor Citrullinated proteins Histones	In situ B cell differentiation Ongoing CSR Autoantibody secretion
Sjögren syndrome	Salivary glands Lacrimal glands	SSA/Ro SSB/La	In situ B cell differentiation Ongoing CSR Autoantibody secretion
Multiple Sclerosis	CNS	Myelin and other neuronal antigens proposed	Unknown
Diabetes	Pancreatic islet parenchyma	Plasma cell anti-insulin Reactivity observed	Unknown
Hashimoto ‘s thyroiditis	Thyroid gland	Thyroglobulin Thyroperoxidase	Unknown
SLE	Tubulointersti tial of the kidneys		Immunoglobulin repertoire analysis reveals B cell clonal expansion and ongoing somatic hypermutation
Cancer			
Primary cancers	Lung Breast Colon	TAA	In situ antigen-driven B cell proliferation, somatic hypermutation and affinity maturation
Infection			
HCV	Liver	Unknown	B cell clonal expansion Elevated CXCL13 expression systemically and locally
Helicobacter pylori	Gastric mucosa	Unknown	CXCR5-CXCL13 dependent Local T cell priming and support T helper 17 cells IgG and IgA Immune response
Mycobacterium tuberculosis	Lung	Unknown	Priming of antigen specific Thelper 1 cells CCL19 and CCL21 expression Accumulation of T follicular helper like cells
Environmental			
Cigarette smoke	Lung	Unknown	Dependent on LTαβ and LTβR CXCL13 and CCL19 expression

Table 1.4. Human chronic inflammatory diseases with ectopic lymphoid structures. This table has been readapted from *“Ectopic lymphoid-like structures in infection, cancer and autoimmunity, Costantino Pitzalis, Gareth W. Jones, Michele Bombardieri1 and Simon A. Jones, 2014”*

CCL, CC-chemokine ligand; CNS, central nervous system; CSR, class-switch recombination; CXCL13, CXC-chemokine ligand 13; CXCR5, CXC-chemokine receptor 5; EGC, ectopic germinal center; HCV, hepatitis C virus; LT, lymphotoxin; LTβR, lymphotoxin-β receptor; SLE, systemic lupus erythematosus; SSA/Ro, Sjögren’s syndrome antigen A (ribonucleoprotein autoantigen); SSB/La, Sjögren’s syndrome antigen B (autoantigen La); TAA, tumour-associated antigen.

Secondary and tertiary lymphoid organs:

In the human body the immune response is mainly orchestrated by complex and organized structures which are classified into primary and secondary lymphoid organs. Primary lymphoid organs comprehend the thymus and the bone marrow, and their main function is to produce and select naïve lymphocytes (B and T) with high diversity in their receptors. Conversely, secondary lymphoid organs (SLO) comprehend lymph nodes, spleen and very specific tissue in the mucosa such as the Peyer's patches.

The main role of these organs is it to allow recirculation of white blood cells and to regulate the encounter between the antigen (trough APC) and the final effectors of the immune system. In order to perform these activities, SLOs are characterised by a fine architecture which allows precise immune cells trafficking and compartmentalization and are anatomically segregated thanks to the capsule. Inside, this system is sustained by a network of stromal cells and crossed by an intricated and specialised vascular and canalicular system, including the afferent and efferent lymphatic system, and specialised blood vessels called the high endothelial venules (HEV) [70-73]. Trough HEVs, naïve lymphocytes expressing L-selectin can access the lymph node from the blood and localize in specific areas: B cells localize into the B cell zones (defined as follicles) in the cortex region - closer to the capsule - while the T cells localise into the T cells zones, (also defined as paracortex), which is more central to the organ. In normal lymphoid conditions, B cells follicles tend to appear close to follicular dendritic cells (FDC) and HEVs and to be surrounded by T cells. Follicles are then divided into germinal centres (GC), whereas - during an immune response – further compartmentalization can be identified: upon activation naïve B cells enter the germinal center and start to proliferate, while undergoing cycles of SHM (dark zone); these centroblast can then move to the light zone, becoming centrocytes, which are cells in a state of active apoptosis competing for survival signals. The highest the affinity induced during the

SHM, the strongest the binding to FDC and T follicular Helper cells (TFH) and the chances for a specific cell to be rescued from the anergetic or apoptotic program. Therefore, the main TFH role is to help the selection of affinity matured B cells clones. Also, these two areas are surrounded by the mantle zone, an enriched T cells area.

More in particular, this peculiar segregation is possible thanks to differentiated gradients of specific cytokines and chemokines produced by the stromal cells network. For instance: T cell who express CCR7 binds CCL19/CCL21; centroblast in the dark zone express high level of CXCR4 which binds CXCL12 produced by reticular cells, while centrocytes in the light zone express CXCR5, which binds CXCL13 produced by the FDCs [74].

TLOs can appear during injury or inflammation and resolve at the end of the response. Nevertheless chronic inflammation can stabilize these structures, allowing the acquisition of SLO features [75]. SLOs and TLOs share similarities in term of segregation of T/B cells areas, and for the appearance of vascular structure analogues to HEV. B cells are divided in the same different areas present in the germinal center where they proliferate, usually close to CD21+ FDC. Plasmablast and plasmacells can accumulate around the T and B areas. A major difference between SLOs and TLOs is the absence of the surrounding capsules. The main function of the capsule is to allow antigen presentation in an isolate environment, while in TLOs the immune cells are directly exposed to all the external signal of the inflamed tissue. These alterations might lead to abnormal T and B cells activation and presentation of autoreactive antigens (such as citrulline). Another main difference is that molecules that are usually strictly produced and regulate only during development in SLOs, are again expressed and activated *de novo* in TLOs. The most important are lymphotoxin-alpha (LT-α) and lymphotoxin beta (LT-β) , produced by the B cells, which usually stimulate the expression of other fundamental cytokines crucial for organization and compartmentalization of the EGCs (e.g. CXCL12, CXCL13, CC19 and CCL21) [76]. A peculiar subset of

activated synovial fibroblasts has been recently described as one of the major sources of CXCL13 inside the inflamed joint [77]. An example of a synovial TLO can be visualized in figure 1.2.

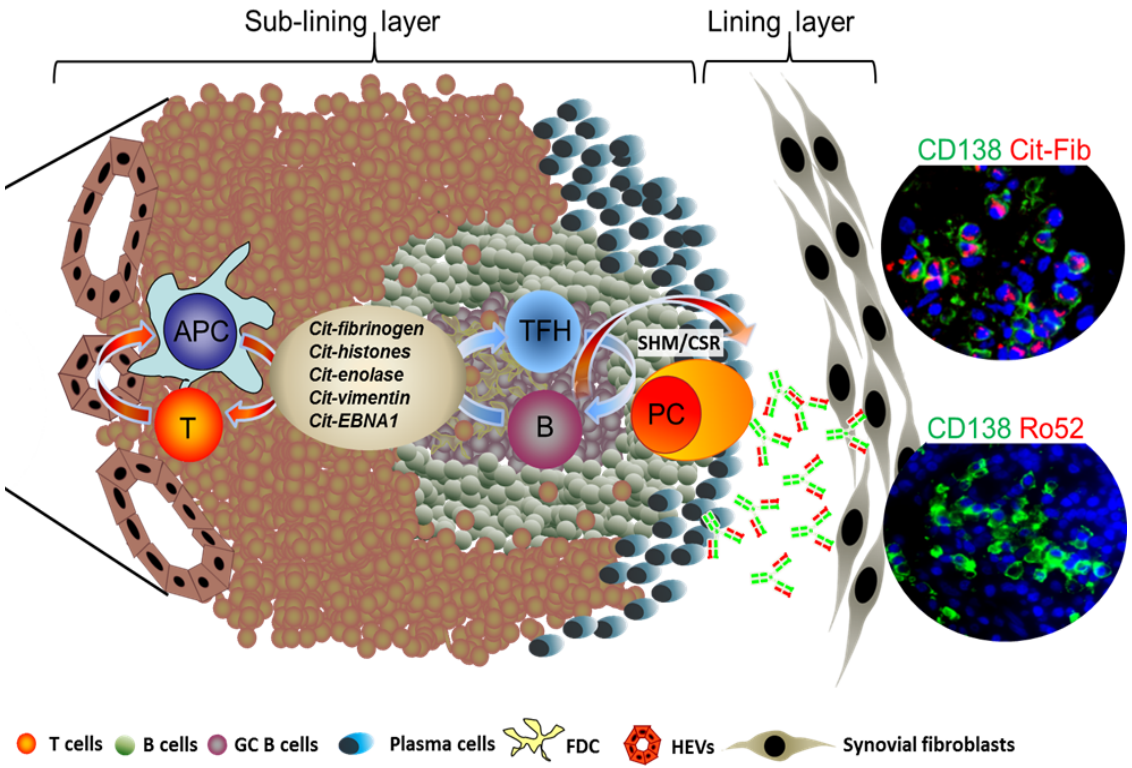

Sub-lining layer
Lining layer
APC
T
Cit-fibrinogen
Cit-histones
Cit-enolase
Cit-vimentin
Cit-EBNA1
TFH
B
SHM/CSR
PC
CD138 Cit-Fib
CD138 Ro52
T cells
B cells
GC B cells
Plasma cells
FDC
HEVs
Synovial fibroblasts

Figure 1.2. Synovial TLO (previous page). Schematic representation of a synovial TLS showing full feature as a secondary lymphoid organ. T cells and B cells are segregated in different areas; cluster of germinal center B cells in the B cell rich area; network of follicular dendritic cells (FDC) derived by stromal cells, who support ectopic GC development; lining layer formed by synovial fibroblasts; development of high endothelial venules (HEVs) at the T rich area; localization of B cells and after activation GC B cells and final differentiation into hypermutated and class switched plasmacells. In autoimmune diseases plasmacells can then produce a consistent titer of autoreactive antibodies, such as ACPA and anti-RO, as demonstrated by immunofluorescence showing colocalization of this cells (CD138) with citrullinated antigens and RO52. (T Cells (T); Antigen presenting cells (APC); B cells (B), T follicular helper, Plasmacells (PC), Somatic Hypermutation (SHM), Class witching Recombination (CSR)) [75] (CC BY-SA 3.0).

B cells development and affinity maturation in normal conditions and in RA

Hematopoietic stem cells (HSC) in the bone marrow are precursor for both T and B cells. While progenitors of the T cells lineage migrate to thymus to finish to complete maturation, B cells development takes place almost entirely in the bone marrow, where niches created by stromal cells provide growth factors, cytokines and survival signals. Then, these signals allow the B cells precursor to survive and further differentiate, going through different stages (pro-B, pre-B and then, finally, immature B cells) each characterised by specific surface markers and by continuous rearrangement of the immunoglobulin (Ig) heavy (H) and light (L). Once this primary rearrangement is completed, the pre-B cell receptor (pre-BCR) is expressed for the first time. The pre-BCR is formed by the μ chain and, after another rearrangement of the Ig L genes, the B cells evolve into IgM^+ immature B cells. The last step in this is the final maturation, where the cells become IgM^+/IgD^+ mature. There is a very refined system of checkpoints and selection in this process, but defects in the maturation pathway can lead to several issues, such as autoimmunity, primary immunodeficiencies or cancer.

Once the resting mature IgM^+/IgD^+ cells are present in follicles inside the SLO, they can become activated by process orchestrated by T helper cells and APCs.

In GC, B cells undergo cycles of SHM of the immunoglobulin variable heavy and light (VH and VL) genes in order to achieve affinity maturation and class-switch recombination (CSR) of the immunoglobulin (Ig) genes and then finally differentiate into memory cells or plasmacells. The latter can then produce a higher titre of autoreactive antibodies (Fig. 1.2). It is important to mention that each of these steps previously mentioned it is defined but expression of specific markers, which are now very well characterized (Fig. 1.3).

In this context, the enzyme activation-induced cytidine deaminase (AID) has been proven to initiate SHM by introducing single point mutation in WRC (W=A/T, A/G) motifs, or so-called mutational hotspot [78]. This allows the B cells, through several cycles of proliferation, to acquire new mutations, and trough this mechanism to generate an incredible diversity. Also, for the IgM^+/IgD^+ to switch for their final function as effector into IgG, IgA and IgE. CSR can take place inside or outside the germinal center and a recent work from Roco, Vinuesa et al revised previously held assumptions regarding CSR. Using an immunization model in which mice have received transferred B cells specific for hen egg Lysozime (HEL), they found that CSR predominantly happen even before the GC or any kind of extrafollicular differentiation [79].

There is now conclusive evidence that B cells plays a fundamental role in the RA synovium [75]. Indeed, my group demonstrated that B cells actively undergo CSR within RA synovium. Humby et al identified the presence of circular transcript (obtained during CSR from IgM to IgG) in AID+ CD21+ RA synovium, supporting the presence of autoreactive plasmacells and high affinity IgG autoantibodies against several citrullinated antigens.

Crucial experiments by different groups have demonstrated that V gene repertoire of B cells obtained from the synovium – both from RNA extracts and microdissection – frequently display on oligoclonal repertoire with highly mutated V regions, which is usually a hint for local antigen-driven GC reaction. The direct evidence that *in situ* diversification in the RA synovium may allow B cells proliferation has been provided by my group, showing high AID level expression in B cells aggregates in close contact with FDC inside the synovial environment [80]. This expression seems to be present in some cases also without FDC.

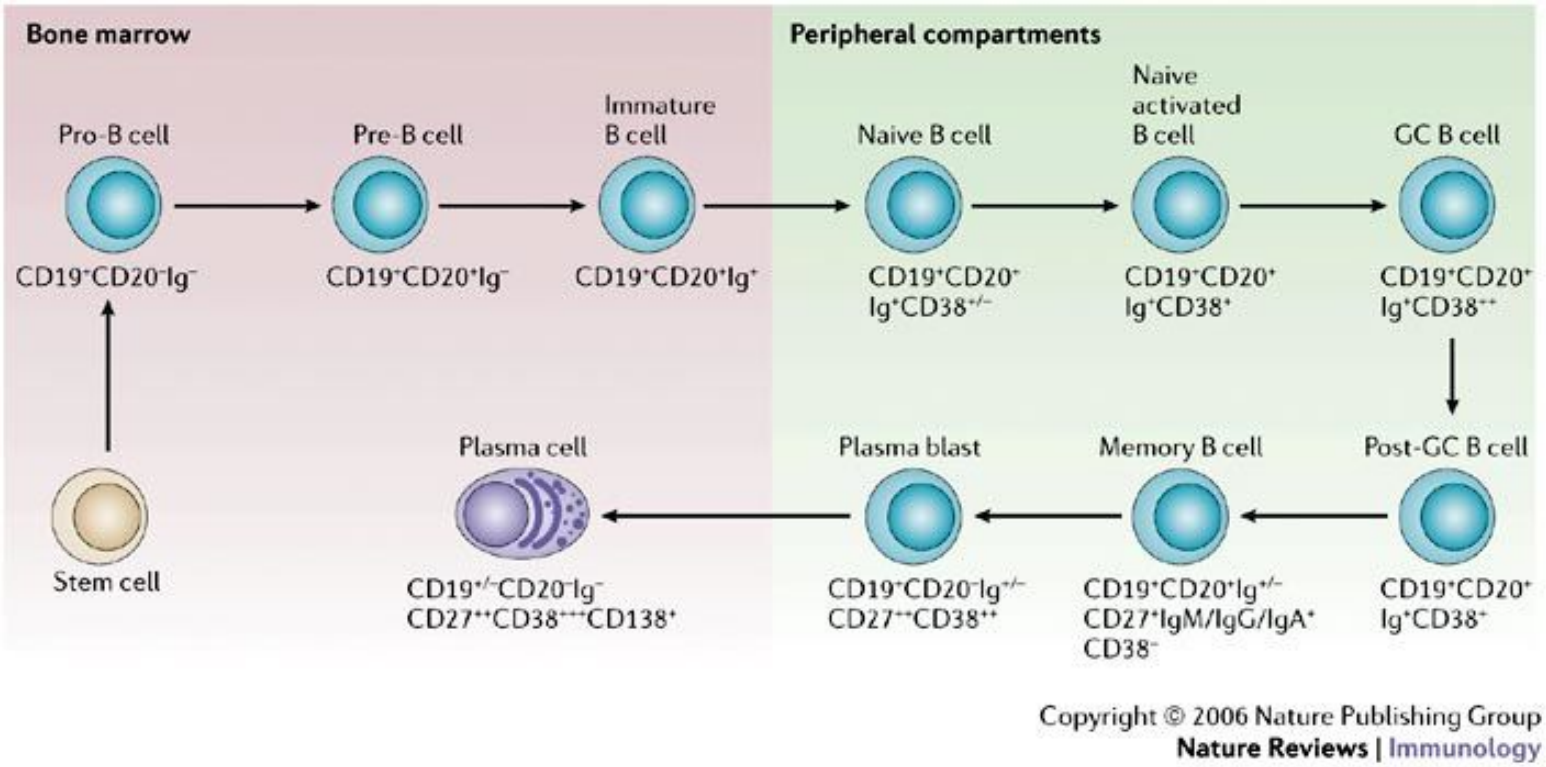


Figure 1.3. B cell development and surface markers expression. Here the expression of different surface markers during B cell development is shown (from *"B-cell targeting in rheumatoid arthritis and other autoimmune diseases, Edwards, J.C. and G. Cambridge, Nat Rev Immunol, 2006. 6(5): p. 394-403"*) [81].

Autoantibodies in RA: RF and ACPA

Many autoantibodies have been connected to RA in the last 70 years, including RF, ACPA and, more recently, anti-carbamylated protein antibodies (anti-CarP) and anti-acetylated protein antibodies. Accordingly, to the presence of RF and/or ACPA, RA patients can be divided and stratified into seropositive and seronegative. The formation of these autoantibodies has been connected with both environmental (smoking) or genetic (shared epitope) factors. Interestingly, in some cases these autoantibodies can be detected many years prior to the onset of the disease, and this might suggest some sort of etiological trigger in which the first autoantibodies appear in predisposed hosts years before an established inflammatory arthritis [82]. Focusing on the antigens targeted by these autoreactive autoantibodies a vast spectrum of elements has been discovered, including stress protein, enzymes, nuclear proteins, cartilage components and, most importantly, several citrullinated antigens (Table1.5). This demonstrates that RA is not characterised by response to a single antigen but by a mixture of different autoreactive elements - connected both to B and T cells – and this mixture may vary during the course of the disease and differ from patient to patient [83].

First evidence that serum from RA patients contained some sort of biological marker activating the agglutination of sheep red corpuscles, thus characterized by autoreactivity, was provided in a pivotal study in 1940 by Norwegian Dr Erik Waaler [84]. Dr Waaler first described the RF which was later redescribed in 1948 by Dr Harry Rose, who also developed an accurate diagnostic test. Nowadays this test is still referred as the Waaler-Rose Test and allowed the discovery that RF is directed to the FC region of IgG [85]. Usually RF manifest as an IgM , but also IgG and IgA have been described [86].

RF is considered to be high in serum when above 20 IU/mL and this usually occurs in 80% of RA cases. Higher levels of RF are clinically connected to a greater probability of destructive articular disease. RF has low specificity for RA being present in up to 50% of Sjögren's syndrome patients and in other autoimmune and infectious diseases. Intriguingly, RF it is also found in Epstein–Barr virus or Parvovirus infection and in 5 to 10% of healthy persons, especially in elderly people.

Although a direct pathophysiological connection has still been difficult to made (this will be discussed later) there is a clear association between RF and strong active synovitis and joint damage.

Autoantigen	Disease	Tissue Localisation
Filaggrin	RA	Epidermis
Vimentin	RA	Mesenchymal cells
α Enolase	RA	Synovium
Fibrinogen	RA	Serum
Collagen Type I, II	RA	Synovium
Fibronectin	RA	Extracellular matrix
Myelin basic protein	MS	Neuron sheath
Eukaryotic translation initiation factor 4G1	RA	Cytoplasm
F-actin capping protein α-1 subunit	RA	Synovium
Far upstream element-binding proteins (FUSE-BPs)	RA	Nucleus
Histones	FS	Neutrophil Extracellular traps

Table 1.5. List of Citrullinated autoantigens. Readapted from "*Citrullination of autoantigens implicates NETosis in the induction of autoimmunity, Nishant Dwivedi, Marko Radic, Annals of Rheumatic Disease, 2014"[87]*. Felty's syndrome (FS), rheumatoid arthritis (RA)

Another mechanism in which RF can contribute to the disease is by formation of immunocomplexes and complement fixation. The former has been demonstrated in a K/BxN mouse model where the development of an inflammatory arthritis is directly connected to the formation of complexes between autoantibodies and glucose-6-phosphate isomerase [88]. There is evidence that the autoreactivity potential of RF manifests both by self-association and by increasing the arthritogenicity of other autoantibodies, such as ACPA [89, 90].

In 1993, another important discovery was made when Serre et al described filaggrin as a target antigen of RA-specific anti-keratin antibodies (AKAs). Further examination led to a deeper understanding that AKAs – but also other RA-specific autoantibodies known as anti-perinuclear factors (APFs) and anti-Sa antibodies – bind to peptides containing citrulline as common antigenic element, nowadays commonly referred as ACPA [91-93]. Citrullination (or deimination) is a reaction mediated by peptidyl-arginine deiminase (PAD) enzymes, which convert the amino acid arginine into citrulline. This modification occurs post-translationally and usually happen in several normal physiological conditions, including maturation of the epidermis [82].

There is a strong association between the presence of the shared epitope and the level of ACPA in the serum of RA patients, and this has led to speculation that peptides presented by SE-containing alleles might be more frequently citrullinated. Indeed, Hill et al have shown in 2002 that the conversion of arginine to citrulline allows for a high-affinity peptide interaction in the MHC class II molecule when the rheumatoid arthritis-associated HLA-DRB1*0401 is present [94]. This has been further confirmed by a cutting-edge paper in which crystal structure studies have shown that the HLA-DRB1 beta chain antigen preferentially bind citrullinate antigens compared to other alleles that bind both arginine or citrulline [95]. Nevertheless, a final culprit regarding the exact binding-motif of the SE molecules *in vivo* has still been difficult to identify. Interestingly, it has been shown that the SE allele is able to interact with

surface calreticulin (CRT), a molecule with a strong connection with citrullination and arthritis. In several works by Ling and De Almeida it was shown that the interaction between SE and CRT is able to skew the immune response to a Th17 subset and to a reduction of Treg [96, 97]. Furthermore, in my lab, Corsiero et al identified synovial B cell–derived RA-monoclonal antibodies (obtained from ELS+ RA synovium tissue) able *in vitro* to react toward FLS-derived CRT, showing that this might be a mechanism contributing to the local autoimmune response [98].

Recently, strong connections have been added between ACPA production and the microbiome, in particular regarding the association between periodontitis and RA. *P. gingivalis* is the main causative agent for periodontitis and his bacterial PAD might contribute to an increase of citrullinate antigen [99]. In addition, ACPA can bind alpha enolase derived from this bacterium and this cross reaction might produce an inflammatory immune response that can trigger the disease. Nevertheless, citrullination pathway of *P. gingavalis* PAD differs strongly from the human PAD, so further experiments are needed to establish the relevance of bacterial PAD in the breach of tolerance to citrullinated antigens [100].

The pathogenic potential of autoantibodies

While the clinical relevance and diagnostic value of autoantibodies have been an uncontroversial dogma in RA, their direct role in the pathophysiology has been difficult to proof. Nevertheless, in the last 20 years several papers have cast light in this direction, finally connecting these molecules to the inflammatory cascade leading to joint damage. One of the lines of evidence come from the fact that rituximab depletion of B cells greatly improves patients conditions, and this is even more marked in ACPA+ and RF+ patients, indirectly suggesting a major contribution of the autoantibodies in RA [82, 101]. In this section I will discuss some of the possible pathogenic roles of ACPA (Fig 1.4).

The first one is connected to the binding of FC receptors on effector immune cells, normally a physiological interaction between the antibodies and the immune cells to mount protective immune responses. Specifically, immunocomplexes containing citrullinated fibrinogen can stimulate macrophage activation through FcγR engagement leading to increased TNF secretion [102, 103].

The damaging effect of immunocomplexes appears to be increased when in synergy with RF-IgM or RF-IgA, as supported also by epidemiological studies [104]. Furthermore, ACPA can directly bind and activate monocytes on GRP78, a citrullinated surface receptor, triggering NF-kB activation and leading to the production of inflammatory cytokines as well as complement activation [105].

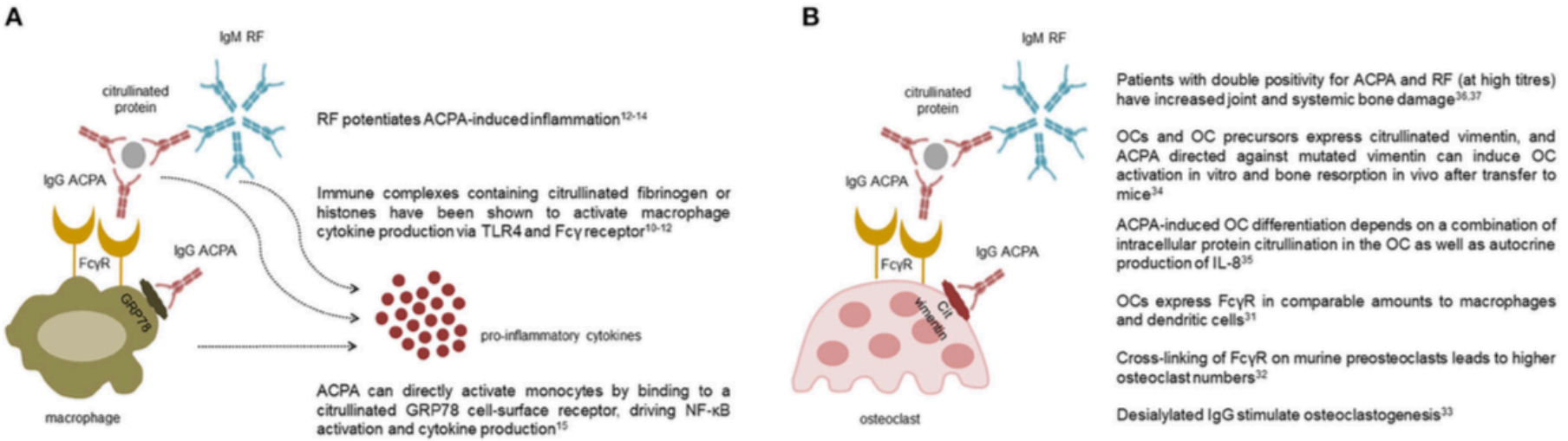


Figure 1.4. Pro-inflammatory and osteoclastogenic roles of autoantibodies in rheumatoid arthritis: From *"Frontiers in Medicine, Bugatti, Manzo, Montecucco and Caporali, 2018: The multiple effects of anti-citrullinated protein antibodies (ACPA) and rheumatoid factor (RF) on macrophages (A) and osteoclasts (B) are shown. Collectively, ACPA are able to stimulate cytokine production by macrophages both through the interaction of their Fc tail with stimulating Fcγ receptors, and through Fab-mediated recognition of membrane citrullinated proteins such as GRP78. Pentameric class M RF potentiates ACPA-induced inflammation likely via activation of the complement cascade (A). ACPA are also able to promote osteoclast differentiation and activation through similar mechanisms mediated by their Fc and Fab portions. Again, clinical studies indicate that RF, especially at high titres, amplifies ACPA-mediate bone damage (B)"*[106] .

The latter can be activated in RA by ACPA through both the classical pathway (initiated by C1q) and the alternative pathway (initiated by C3) but not the lectin pathway [107]. Indeed the synovial fluid of RA patients shows low C1q and C3 together with an increase of cleavage products, indicating a strong complement activation [108].

Another important feature in RA is the augmented generation of neutrophil extracellular traps (NETs) from both circulating and synovial fluid neutrophils. Indeed, both ACPA and RF have shown to strongly influences this mechanism. NETs formation is an elegant physiological event where neutrophils protrude condensed chromatin to trap and kill pathogens. The link between NETs and ACPA production is extremely interesting as during NETtosis, chromatin decondensation requires the hypercitrullination of histone proteins, leading to excessive exposure of citrullinated antigens that may lead to the production of ACPA. Interestingly Pratesi and Migliorini have provided evidence that ACPA can in turn stimulates NETtosis and target NETs antigen, fuelling the ongoing ACPA production, opening the scenario to a process that might continuously renovate itself, leading to a systemic non-resolving response [109]. My group has demonstrated that monoclonal antibodies obtained from synovial fluid B cells of RA ACPA+ patients have a strong reactivity against citrullinated histones produced during NETs formation [110]. Some studies have also connected NETs and RA-FLS, but this will be discussed later. Recent experiments have shown that IgA-RF immune complexes can result in an increased NETosis [111, 112].

RA patients who are RF+ or ACPA+ are usually affected by a more severe form of RA, with enhanced joint destruction [113, 114]. Several works have shown that ACPA are able to interact with FcγR on osteoclasts, both *in vitro* and *in vivo*, and also to induce their differentiation from osteoclast progenitors. In 2016, Krishnamurthy et al have identified a novel IL-8 dependent molecular mechanism that link ACPA to osteoclasts and direct bone erosion, by showing that IL-8 antagonist can rescue bone destruction.

In this context ACPA seem able to bind citrullinated vimentin on osteoclasts surface [22]. As with immunocomplexes, osteoclasts activation - and thus degradation of bone - seems to be increased when RF and ACPA are present in combination, as supported by increased radiographic progression in double positive RA patients [106].
Recent works have also suggested a direct role for ACPA in promoting pain. Wigerblad et al injected murinised monoclonal ACPA in a B10.RIII mice mouse model managing to trigger a directly induced pain response. This mechanism seems to require the binding of autoantibodies to marrow osteoclast precursor, which would in turn lead to the production of the IL-8 murine analogue CXCL1 [115]. This novel work shows how ACPA can directly contribute to arthralgia, one of the most important symptoms in RA.

The role of FLS and their interactions with immune cells

As already mentioned, B cell activation is sustained by a network of FDCs that have the same phenotype and functional capacity as FDCs in lymphoid organs and which are derived from locally differentiating stromal cells, probably FLS [116]. It has long been known that FLS play a fundamental role in the pathogenesis of RA either producing a wide array of pro-inflammatory mediators (cytokines and chemokines) or directly contributing to local cartilage destruction due to their neoplastic-like behaviour in terms of tissue invasion and matrix degradation [117]. It has now been extensively shown that resident joint tissue synovial fibroblasts - but also dermal, gingival and intestinal - can exert immunological functions. Indeed they display several immune characteristics like expression of innate immune receptors, i.e. TLR [118, 119], production of cytokines and chemokines in the presence of pathogens or trauma, ability to process antigens and expression of major histocompatibility complex (MHC) receptors [120]. In a healthy joint, synovial fibroblast are the most

abundant cells, constituting one of the crucial elements in both lining and sublining layer, the latter presenting with a normal thickness of 1-2 cells. They are the major contributor to the lubrication and integrity of the cartilage, thanks to the production of several components of synovial fluid, such as the extracellular matrix (ECM). These mesenchymal cells, also known as type B synoviocytes - to distinguish them from the type A macrophage-like synovial cells - have a clear fibroblast phenotype in term of expression of type IV and V collagen, vimentin and CD90 (Thy-1). Nevertheless, FLS have some unique characteristics compared to other fibroblast lineages, including the expression of Cadherin-11, an adhesion molecule important for homotypic aggregation [121].

This delicate equilibrium is totally altered when a patient is affected with RA, whereby the synovial lining layer presents a marked increase in thickness and immune cells start to infiltrate the sublining, establishing cell-cell contact interactions with FLS.

It has long been debated whether synovial stromal cells are simple bystanders and inert actors during the development of RA or they rather play a critical role in initiating and propagating the disease (the so called "mesenchymal hypothesis" of RA, put forward by Firestein [122]) . A series of studies in the last two decades have cast new light on the crucial importance of fibroblasts in the pathogenesis of RA, a concept which is currently well accepted [123]. One of the most important characteristics of these cells is that, upon proliferation, they become activated and ready to invade and destroy the cartilage tissue. For most of the studies performed worldwide, RA-FLS have been taken from patients undergoing joint replacement or synovial biopsies. However, it is imperative to highlight that long-term culture may affect both phenotypes and genotypes of these cells, as already been discussed by other authors [124].

Focusing on the role as immune cells, it has been showed that FLS express TLR1-7 but very low or basal transcript level of TLR8, -9 and -10 [118]. A summary on the role of each single TLR on RA-FLS has been collected by Ospelt in 2017 [125]. Briefly, while TLR-3 and -7 are sensors for double and single-stranded RNA respectively, TLR-2 (which use TLR-1 and -6 as heterodimers) and TLR-4 are sensors of bacterial activity, detecting the presence of lipoproteins such as LPS. As demonstrated in my lab by Kam and Bombardieri, upon stimulation of TLR-3, RA-FLS are able to produce high amount of B- cell activating factor (BAFF) and A proliferation-inducing ligand (APRIL) [126]. In particular, several endogenous ligands, which are usually increased in an RA joint, have been identified *in vitro* for their ability to activate one or more TLR pathways. Specifically, cells undergoing necrosis can release self-molecules able to trigger TLR-3 signalling [127], while fibrinogen can bind TLR-4, showing an even more potent activation in its citrullinated form. Similarly, free histones, which have been described to bind both TLR-2 and -4, were able to activate TLRs with higher efficiency in their citrullinated form or when presented as immunocomplexes. Furthermore, this result has been linked to the

subsequent presence of autoantibodies, especially ACPA, as described by Hatterer et al [128]. Only a few studies have focused on DNA sensors, such as TLR-9, since 24 hours stimulation of RA-FLS with unmethylated CpG rich DNA has not shown any increase in inflammatory cytokines such as IL-6 or IL-8 [129]. Only recently Carmona-Rivera et al have described the ability of RA-FLS to internalise (NETs)-derived citrullinated histones through a receptor for advanced glycation end products (RAGE) and TLR-9 pathway [130]. Furthermore, they showed that FLS primed with NETs products are able, in a mouse model of RA, to promote ACPA production. Nevertheless, some of these results have been criticised and the presence of TLR-9 on RA-FLS is questionable.

It is important to mention that TLRs are not the only innate immune pathogen recognition receptors (PRR) described in FLS. Also, retinoic acid-inducible gene I (RIG-I) and MDA5 (which bind nucleotides), NOD1 (recognizing bacterial molecules containing D-glutamyl-meso-diaminopimelic acid (iE-DAP) moiety) and NOD2 have been described in these cells [131-134].

For what concern the interaction with other resident as well as infiltrating immune cells, synovial fibroblasts can orchestrate strong interaction with T cells, monocytes, macrophages, endothelial cells and osteoclasts [135-138]. Furthermore, RA-FLS likely have a role in the formation of ELS and B cells survival.

As above mentioned, RA patients can be stratified - following histological analysis - in three different pathotypes: i) the pauci immune or “fibroid pathotype”, characterised by an increase of the fibrotic component and a lower number of immune infiltrates; ii) the diffuse myeloid pathotype, where monocytes and macrophages are the most abundant cells; iii) the lympho-myeloid pathotype, where B an T cells aggregates are predominant and can often organise themselves in [65, 139].

Recent evidence is leading to the idea that RA-FLS might promote and orchestrate formation of ELS trough the production of receptor activator of NF-κB ligand (RANKL), CCL19 and CCL21 [75, 140]. Furthermore, the

high level of APRIL, BAFF - produced upon stimulation of the TLRs - and CXCL12, are able to create the perfect environment for the survival of B cells and plasma cells, including potentially ACPA producing-autoreactive clones.

Beyond all the described indirect effects on other cell types, one of the characteristics which allow RA-FLS to actively contribute to the disease is the invasion and destruction of the cartilage. One of the prerequisites for the attachment of these cells to chondrocytes appears to be the loss of key molecular elements of the cartilage, such as proteoglycan, as shown in a TNF transgenic mouse model by Korb-Pap et al [141]. Taken together these results have raised the hypothesis that activated synovial fibroblasts might be able to migrate through the blood stream from an inflamed joint to a non-inflamed joint, with the subsequent establishment of a new destruction process. Among the adhesion molecules, integrins and syndecans have been implicated in the attachment of fibroblasts to the cartilage matrix, specifically α5β1 integrin which binds fibronectin, a protein markedly elevated in the RA joints [142]. Another peculiarity of RA-FLS is the production of destructive enzymes, which is often stimulated by TNF and IL-1. Among these the most important are matrix metalloproteinases (MMP 1,3 9, 13, 14 and 15), a disintegrin and metalloproteinase with thrombospondin motifs (ADAMTs) and cathepsins B, L, and K [143] .

From a gene regulation point of view, it is important to mention that several studies are bringing novel evidence that the observed pro-inflammatory stable phenotype of RA-FLS is due, at least in part, to epigenetic modifications (DNA methylation, histone deacetylation) and transcription repression by microRNASs [144-146].

Another assumption which has been recently been challenged is the phenotypical stability of RA-FLS once activated. The introduction of fluorescence-activated cell sorting (FACS) for the characterization of FLS - a technique which has been underused with these cells due to intrinsic technical difficulties - together with immunostaining and genotypization,

have lately changed this belief. Indeed, FLS from different joints differ in term of function, interaction and, moreover, gene expression compared with other stromal or hematopoietic cells. This has been shown in particular for Hox genes, a group of genes that controls limb development from the embryonic state, and that are expressed as blueprints of their specific location. Frank-Bertonceli et al demonstrated that FLS from different limbs maintain this blueprint established during embryogenesis in adult life [147].

Furthermore, the synovia itself displays heterogeneity in terms of stromal cell populations. The lining layer stroma have abundant VCAM1, Cadherin 11 and podoplanin, while synoviocytes in the sublining express more CD90 and CD248 [121, 148].

Regarding this aspect one of the most interesting recent studies has been published by Mizoguchi et al in 2018, where they identified in fresh human synovium different subsets of RA-FLS, most of them Podoplanin+ and Cadherin11+ but further distinguished for their expression of CD34 and CD90 [149]. CD34 is a transmembrane phosphoglycoprotein, first identified on hematopoietic stem and progenitor cells but now also known to be express on MSC and other stromal cells, including corneal keratinocytes and epithelial progenitor [150]. The role of CD90, instead, has not yet been fully elucidated but a role in cell-cell and cell-matrix interactions, with implications in inflammation and fibrosis has been suggested [151]. One subset in particular, CD34-CD90+ seems remarkably increased in the sublining area around blood vessels, and this expansion has been connected to disease activity, tissue invasion and infiltration of other immune cells. Also, gene expression profiling analysis of this subset has highlighted a high expression of RANKL and a low expression of osteprotegerin (OPG, a negative modulator of osteoclastogenesis) suggesting their role in bone loss. On the other side, CD34-CD90- fibroblasts- whose ability to promote osteoblastic bone formation trough the expression of BMP-6 has led to their identification as a “protective” subset - seems strongly reduced in the RA synovia. A third subset, CD34+,

appeared to be transcriptionally imprinted as a strong secretory rather than an invasive phenotype, due to the high expression levels of IL-6, C-C Motif Chemokine Ligand 2 (CCL2) and C-X-C motif chemokine 12 (CXCL12), among others.

In addition, Choi et al have identified the expression of supplementary FLS markers in early arthritis patients. These include the fibroblast activation protein (FAP) and the decay-accelerating factor (DAF) or CD55 [152]. FAP is a cell surface type II glycoprotein with ectoenzyme activity and, interestingly appears to be expressed on stromal fibroblasts of 90% of human carcinoma. It has been implicated in the control of fibroblasts growth during development and tissue repair. Conversely, DAF, which is a 70 kDa membrane glycoprotein that attaches to the cell membrane via a glycophosphatidylinositol (GPI) anchor, is widely distributed among hematopoietic and non-hematopoietic cells. On FLS, DAF seems to mediate contacts with CD97 on other immune cells and may be of primary importance in maintaining and amplifying synovial inflammation. It has also been linked to interactions with complement proteins.

This work on fibroblast subsets has been very recently pushed forward by Croft, Barone and Buckley who identified and described the biological behaviour of FLS responsible for mediating either inflammation or tissue damage in arthritis. In an elegant *in vivo* work, they have shown that deletion of fibroblast activation protein-α (FAPα)+ fibroblasts suppressed both inflammation and bone erosions of resolving arthritis. Specifically, using single-cell transcriptional analysis, they identified two main fibroblast subsets among the FAPα+ population: FAPα+THY1+, resident in the synovial sub-lining and with a role as immune effector; FAPα+THY1−, a more aggressive and invasive phenotype resident in the synovial lining layer. Furthermore, *in vivo* cell transfer experiments in murine joints have shown very different behaviour of these two subsets. When FAPα+THY1− fibroblasts were transferred, they acted as mediators of bone and cartilage destruction with small effect on inflammation; on the opposite the transfer of FAPα+ THY1+ fibroblasts resulted in an

orchestrated and persistent inflammatory arthritis, with fewer effect on bone and cartilage[77].

From a metabolic point of view a very inclusive review on FLS has been written and published by Falconer in 2018, highlighting a sort of metabolic imprinting of synovial cells - due probably to epigenetics changes - promoted by the inflammatory microenvironment, which is long lasting in in vitro cultures [153]. Several of these metabolic studies have focused on the mitochondrial metabolism and dynamics in order to obtain a complete and more adequate picture of the synovia, but further studies are needed to fully clarify the relationship between metabolic changes and the phenotypic and functional properties of RA-FLS.

The Epstein-Barr Virus

Discovery and History of EBV

It has been more than 50 years since the EBV, the first tumour-inducing virus, was discovered in London. In 1964, a group of scientists published on the Lancet the paper "Virus Particles in Cultured Lymphoblast from Burkitt's Lymphoma" which will radically change the cancer field, and among them, there are two young researchers Anthony Epstein and Yvonne Barr who will later name the virus.

In order to the get the fascination of this discovery a step back in time and space is needed.

Dr Denis Parson Burkitt was working in equatorial Africa - and specifically in Uganda - when, in 1958, he started to investigate a form of cancer which was very common in young kids in that area. It was characterised by painful face or neck swelling and was rapidly fatal if left untreated. He later discovered that the tumour was caused by an uncontrolled and aggressive proliferation of white blood cells. Unfortunately, he had no idea what was causing the tumour but, thanks to some epidemiological studies, he noticed the disease was following a similar pattern to malaria, since it was localized in areas with hot weather and heavy rainfall. This would have been the first hint the Burkitt's lymphoma - later described after him – was probably transmitted by some sort of microorganism.

A few years later, in 1961, Burkitt was presenting in London a lecture entitled 'The commonest children's cancer in tropical Africa: a hitherto unrecognized syndrome', containing most of his studies and hypothesis and, luckily, Epstein was among the audience. It was thanks to his prior knowledge of both Rous Sarcoma Virus (RSV) – a transmissible agent causing tumour in chicken – and the further advantage of having learned a brand new technology for the time – electron microscopy - in the Rockefeller Institute laboratory of George Pallade, that Epstein was able to advance the idea that an oncogenic virus depending on some sort of vector

would be involved. This simple question would have started a long-standing collaboration between the two scientists but, most importantly, it would set the dawn of human tumour virology.

Unfortunately, for 3 long years, every kind of attempt of retrieving Burkitt's sample from Africa and to culture the white blood cells from the tissues had miserably failed and Epstein and Barr – at the time working with Burt Achong – where left with nothing more than frustration. Quoting one of the most elegant review written on EBV by Young, Yap and Murray, *"It is here that serendipity played a part: a delayed and rerouted flight from Kampala in December 1963 resulted in a BL biopsy reaching the London laboratory later than expected. The transit medium in which the BL biopsy was shipped was unusually cloudy but, rather than the suspected contamination with bacteria, microscopic examination revealed viable tumour cells floating free of the main lymphoma mass. These cells grew when suspended in fresh culture medium, requiring the culture to be split into new culture bottles. And so, the first cell line from a human lymphoma was established, and early in 1964, Epstein prepared these BL cells (now named the EB1 cell line after Epstein and Barr) for electron microscopy. He was exhilarated to observe unequivocal virus particles in a cultured BL cell and it was clear from the morphology of the virus particles that this was a herpesvirus. Unlike the three other human herpesviruses known at that time, this virus was inert as it did not induce the characteristic lytic infection usually associated with herpesvirus replication. These observations were published in a Lancet paper that became a citation classic in 1979"*[154, 155]. For a timeline of the major milestone in EBV research check figure.1.5.

Classification of EBV

The EBV - also called human herpes virus 4 (HHV-4) - is classified within the gammaherpevirus subfamily and has to be considered as the prototype of the genus *Lymphocryptovirus*. In human two major EBV types have been detected, relatively EBV-1 and EBV-2. Formally they differ in the

EBNA expression sequence (EBNA-2, EBNA-3A, EBNA-3B, and EBNA-3C) [156]. These two subtypes - also referred as type A and type B - differ in the way they are able to immortalize cells lines, with EBV-2 subtype having less immortalizing efficiency than EBV1, probably due to mutation in EBNA-2 [157, 158].

Nevertheless, in 2019 Zanella et al have published a newer and more specific classification based on phylogenomic and population analysis [159]. In their work, they criticised the old classification claiming that only small subsets of genes have been used to sort out strains and this have led to failure in identifying all the genomic variability and recombinant region present in EBV genome. In order to avoid misinterpretation and to minimize recombination signals - which appears to be 2.5-times stronger that the mutation events - they used a hierarchical bayesian analysis of population structure (hierBAPS), which resulted in the differentiation of 12 EBV populations. These populations are mostly related among them due to geographic location and tumour development.

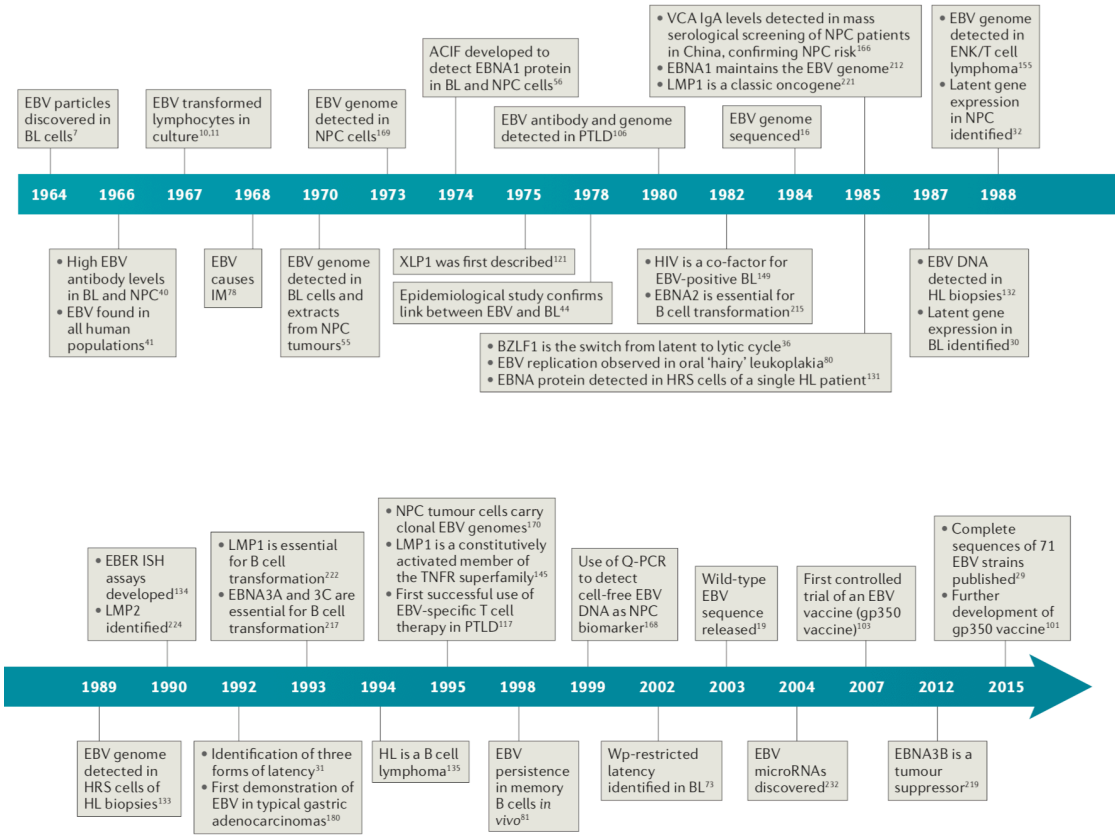


Figure 1.5. From *"Nature Perspectives, Young, Yap and Murray, 2016: Timeline showing the major milestones in EBV research. ACIF, anti-complement immunofluorescence; BL, Burkitt lymphoma; EBER ISH, Epstein–Barr virus-encoded RNA in situ hybridization; EBNA, Epstein–Barr nuclear antigen; EBV, Epstein–Barr virus; ENK, extranodal natural killer; HIV, human immunodeficiency virus; HL, Hodgkin lymphoma; HRS, Hodgkin–Reed–Sternberg; IgA, immunoglobulin A; IM, infectious mononucleosis; LMP, latent membrane protein; NPC, nasopharyngeal carcinoma; PTLD, post-transplant lymphoproliferative disease; Q-PCR, quantitative PCR; TNFR, tumour necrosis factor receptor; VCA, viral capsid antigen; XLP1, X-linked lymphoproliferative disease type 1"* [155].

EBV Pathologies and Epidemiology

EBV is one of the most common viruses among the general population. Children in developing countries usually acquire the infection during the first five years of life, while in developed countries this event is usually delayed to adolescence. The presence of the virus inside the B cells is usually asymptomatic but, in some cases, can lead to an infectious mononucleosis which usually resolves in 2 weeks. The virus then survives in patients with an asymptomatic infection and sign of previous infection can be usually found in 90% of adults [160].

The virus is strongly associated with several other diseases, in particular some types of cancer, such as Hodgkin's lymphoma, Burkitt's lymphoma, gastric cancer and nasopharyngeal carcinoma (for a detailed list of EBV-associated diseases check table 1.6) [161]. However, in the last 30 years the strong lymphotropic behaviour of EBV have gained interest in its ability to modulate immune responses with mounting evidence regarding the ability of EBV to induce the breach of self-immunological tolerance and the onset of autoimmunity [162, 163].

Type of disease				**Disease**
Nonmalignant disease				• Infectious mononucleosis • Chronic active infection • Oral hairy leukoplakia
Malignant disease	Immunocompromised host	B cell malignancies	Acquired immunodeficiency	• AIDS-associated B cell lymphomas • Post- transplantation lymphoproliferative disorder • Lymphomatoid granulomatosis • Methotrexate-associated B cell Lymphoma
			Congenital Immunodeficiency	• Severe combined immunodeficiency-associated B cell lymphomas • Wiskott-Aldrich syndrome-associated B cell lymphomas • X-linked lymphoproliferative disorder-associated B cells lymphomas
		Mesenchymal malignancies		• Leiomyosarcoma
	Immunocompetent host	B cell malignancies		• Burkitt lymphoma
				• Classical Hodgkin lymphoma
		T cell malignancies		• Extranodal NK/T cell lymphoma, nasal type • Virus-associated hemophagocytic syndrome T cell lymphomas
		Epithelial cell malignancies		• Nonglandular nasopharyngeal carcinoma • Lymphoepithelioma-like carcinoma (salivary, thymus, lungs, stomach) • Breast carcinoma • Hepatocellular carcinoma
		Mesenchymal malignancies		• Follicular dendritic cell sarcoma

Table1.6. Epstein Barr Virus-associated diseases. Readapted from *"The Annual Review of Pathology: Mechanisms of Disease Spectrum of Epstein-Barr Virus-associated Diseases, J.L.Kutok and F.Wang, 2006 ": List of Epstein Barr-associated diseases"*. The list does not include Multiple Sclerosis [161].

EBV Structure and Genome

EBV has a diameter of approximately 122-180 nm and its main structure is composed of i) a double helix of DNA containing 172'000 base pairs and around 85 genes [164]; ii) a protein nucleocapsid protecting the DNA; iii) a protein tegument surrounding the nucleocapsid which in turn is surrounded by a structure - called the envelope - containing both glycoproteins, essential for infection, and lipids [165] (fig. 1.6 and 1.7).

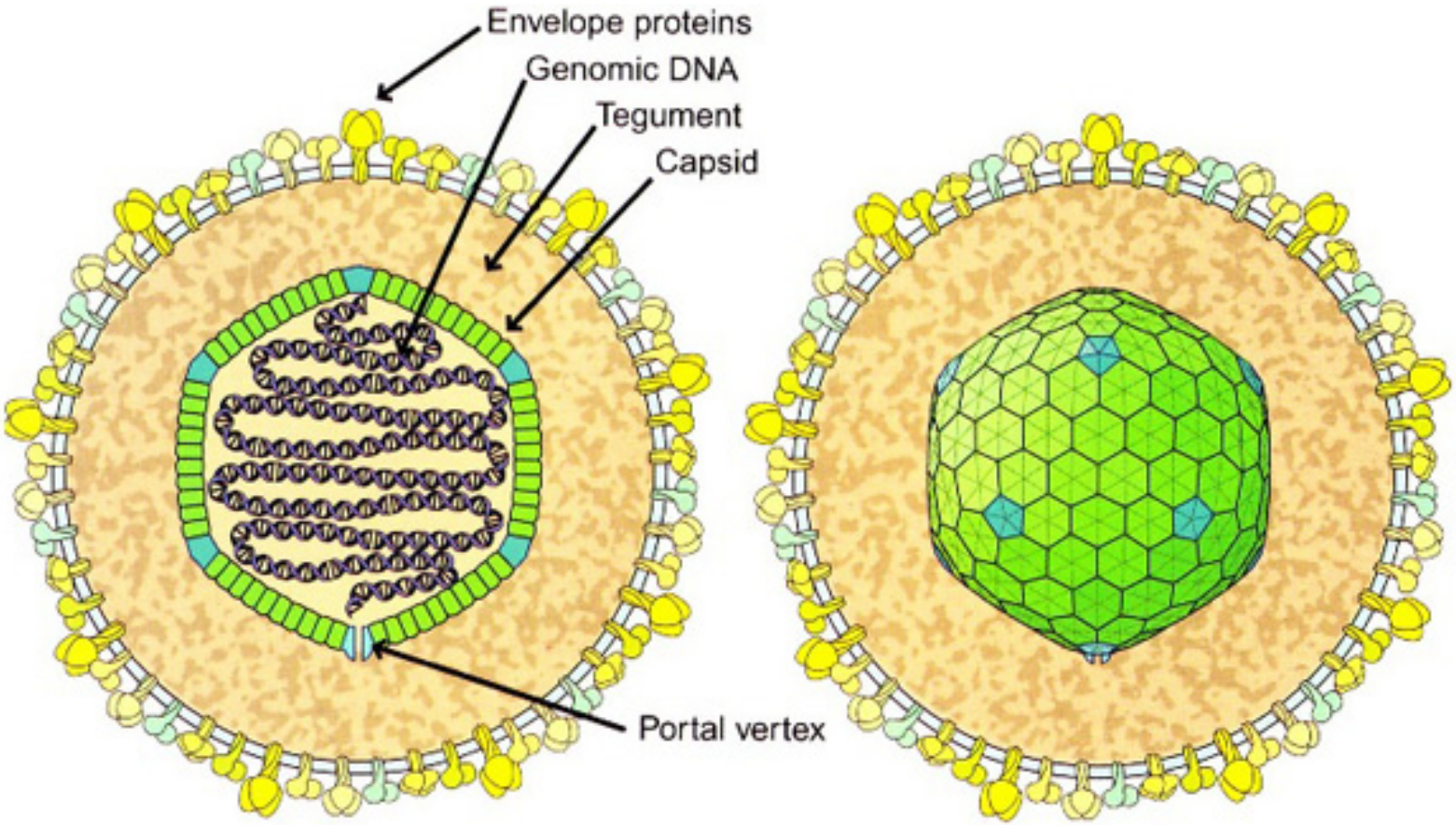


Figure 1.6. Epstein Barr Virus Structure. From "The role of Epstein-Barr viruses in Burkitt' Lymphoma, Mary Frazier Greene, 2014". The structure of the EBV is characterised by the presence of an envelope covered in proteins, a viral tegument, and a viral capsid containing DNA.

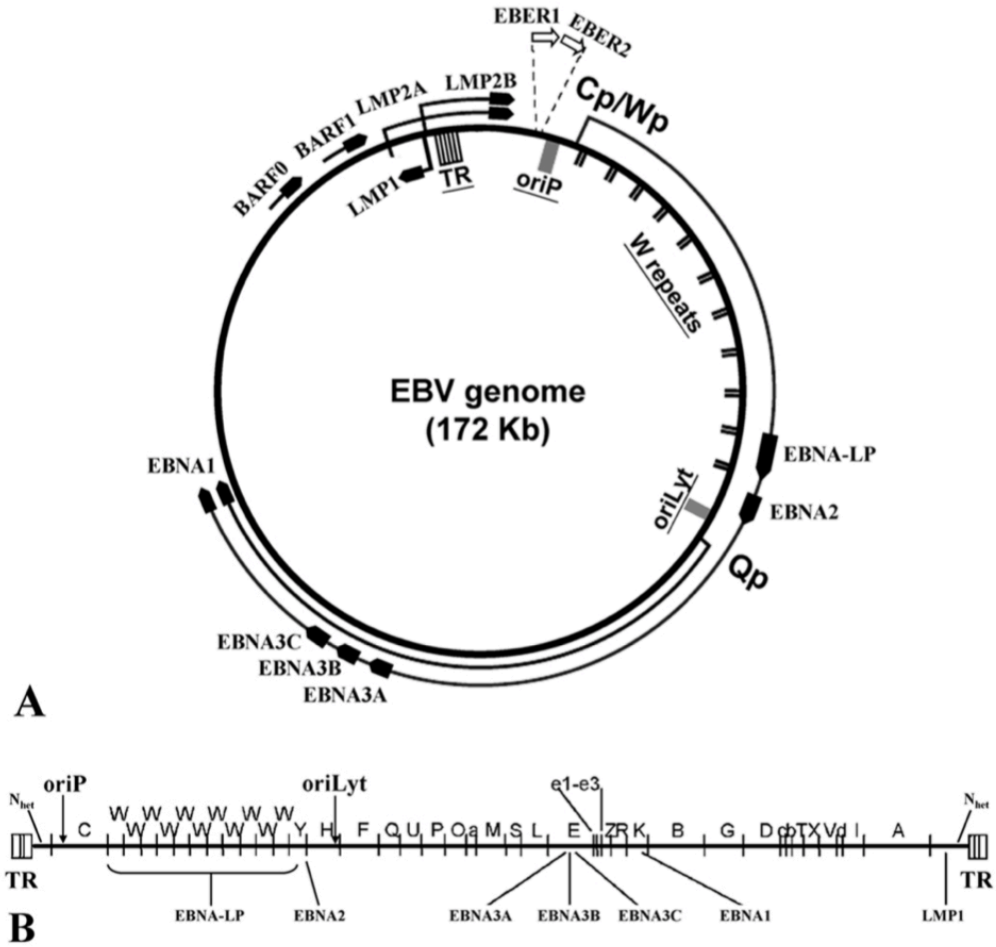
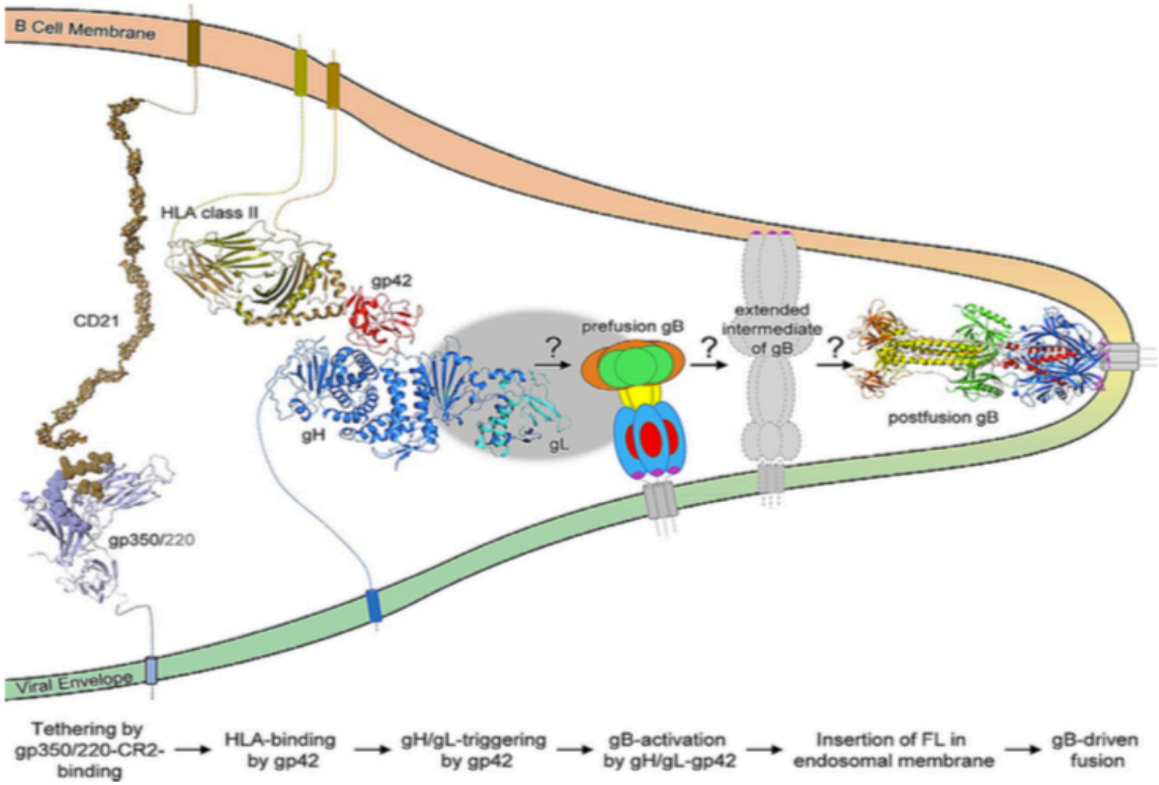


Figure 1.7. Epstein Barr Virus genome. EBV can have both a double-stranded DNA episome (A) or a linear form (B), with the origins of plasmid replication (oriP) and lytic replication (oriLyt) indicated. The Epstein-Barr nuclear antigen leader protein (EBNA-LP), crucial for the latent stage is transcribed from the BAMHI W, a region of variable repeated exon. From "*Tao, Q., et al., Epstein-Barr virus (EBV) and its associated human cancers-genetics, epigenetics, pathobiology and novel therapeutics. Front Biosci, 2006.* ***11****: p. 2672-713*" [166].

EBV Cell entry

One of the peculiarities of the virus is its lymphotropic tropism toward B cells. Nowadays it is known that the virus has several cell entry mechanisms but the preferentially adopted one (and most studied) is the binding to CD21 - also known as CR2 complement receptor – or CD35 by the viral envelope protein gp350/220 [167]. EBV can also infect other cells types, such as mesenchymal cells, T cells and epithelial cells. The latter can be targeted mainly via gHgL which - together with gp42 and gB - forms a viral glycoprotein complex that plays a fundamental role in term of membrane fusion.

Focusing on the fusion machinery that allows to entry into the cells, it is important to mention that - as other herpesvirus – EBV adopts three to six envelope proteins to perform all the operations needed, binding to the host receptor, activate the fusogen and the virus-driven fusion process [168]. After the binding of gp350/220 to the CD21, the virion enters the host cells by endocytosis, and this event greatly increases the efficiency of infection. Nevertheless, this step is not essential, and the virus is able to find alternative ways. Indeed, Nemerow and Cooper demonstrated in 1984 that B lymphoblastoid cell lines are infected by a direct fusion of gp42 (part of the gHgL complex) to the B cells specific HLA class II receptor [168, 169] (Fig.1.8).


B Cell Membrane
HLA class II
gp42
CD21
prefusion gB
extended intermediate of gB
postfusion gB
gH
gL
gp350/220
Viral Envelope
Tethering by gp350/220-CR2-binding
HLA-binding by gp42
gH/gL-triggering by gp42
gB-activation by gH/gL-gp42
Insertion of FL in endosomal membrane
gB-driven fusion

(From previous page) Figure 1.8. Cartoon showing the endocytosis entry pathway of EBV into B-Cell: From Advances in Virus research, Mohl, Chen and Longnecker, 2019: After gp350/220 binds the CD21, gp42 can interact with the MHCII molecule. This leads to conformational changes in the gp42 that may activate gp42 itself or gH/gL allowing the final binding of gB. gB will then rearrange and finally fuse with the host cell membrane. For colour map check reference. From *"Mohl, B.S., J. Chen, and R. Longnecker, Gammaherpesvirus entry and fusion: A tale how two human pathogenic viruses enter their host cells. Adv Virus Res, 2019. 104: p. 313-343"* [168]..

This peculiar alternative entry mechanism has been connected by Trier et al directly to the RA-shared epitope and thanks to their findings on the MHCII crystal structure they suggest that "*EBV is able to induce the onset of RA in predisposed SE-positive individuals, by promoting entry of B-cells through direct contact between SE and gp42 in the entry complex*" [170] (Fig. 1.9).

Once the virus aggregates within the membranes it can be internalized inside cytoplasmic vesicles, where it can survive [169, 171].

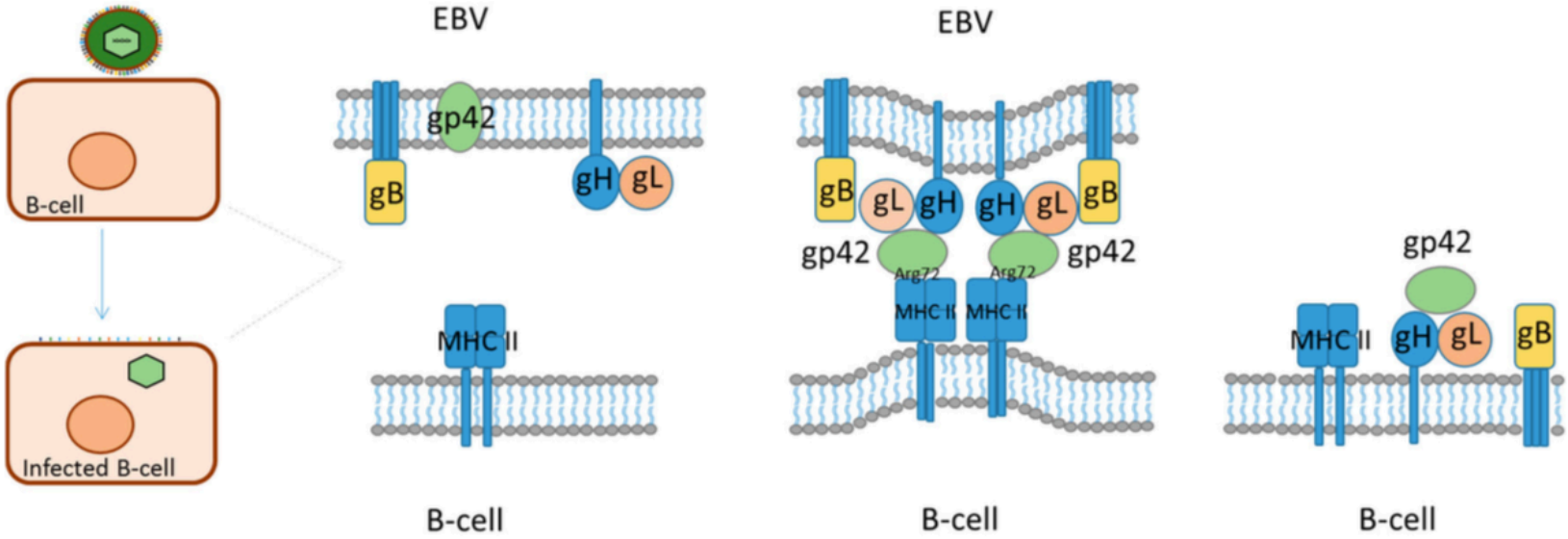


Figure 1.9. Cartoon Showing the EBV fusion Model for gp42/MHCII: From International Journal of Molecular Sciences, Trier et al, 2018. Once gp42 is active it will interact with gh/gL and the entire complex interact with gB. Finally, the gp42 will interact with the β1 domain of the MHCII, allowing the membrane to fuse and the virus to enter the host. From *"Trier, N., et al., Human MHC-II with Shared Epitope Motifs Are Optimal Epstein-Barr Virus Glycoprotein 42 Ligands-Relation to Rheumatoid Arthritis. Int J Mol Sci, 2018. 19(1)"* [170].

EBV Primary infection and persistence: from the germinal center to the memory compartment

The mucosal tissue of the oropharynx is usually where the EBV primary infection takes place. However, the first cell to be targeted by the virus is still a matter of discussion, being either the epithelial cells or the B cells. Several theories have bene proposed in the last 40 years, with the epithelial model as the most accepted: viral particles - present in the saliva of infected individual – might access the oral cavity of the host, and specifically the lymphoid tissue. Lytic replication would take place inside the epithelial cells which then, in turn, would infect the B lymphocytes. Several observations are pointing in this direction: EBV replication has been observed in normal tongue epithelium, suggesting that the tongue is the source for EBV in saliva [172]; patients with IM usually presents EBV in desquamate epithelial cells but also viral DNA, mRNA and proteins [173, 174].

Nevertheless, a discreet amount of data has shown that EBV can directly gain access and infect B cells trough crypt structures of the tonsillar lymphoepithelium [175].

Despite all these controversies it is quite obvious that the B cells are the main target of the virus since - during the maximum disease activity in IM - the number of EBV infected cells can account for almost 1% of the circulating B Lymphocytes [176]. After this stage EBV mostly survives in a silent form inside B cells (named latency phase) with occasional cycles of replication (named lytic phase) (fig.1.10). The infectious viruses are produced during this lytic phase, with approximately 80 viral proteins expressed, including viral capsid antigens, structural proteins and transcription factors. The switch between the latency and lytic stage is possible thanks to BZLF1 (also known as Zta, EB1, associated with its product gene ZEBRA) and BRLF1(associated with Rta).

The alternative pathway for the virus consists in the already mentioned latent phase, a tool for EBV to become “stealth” inside the host’s blood and

survive for years. During the latency phase EBV infected B cells can show four different latent programs, each of them with a distinct EBV gene expression profile, including EBV EBNAs, latent membrane protein 1 (LMP-1) and LMP2A [177]. Importantly, these two last proteins are involved in a mechanism, called transformation, that can actually immortalize B cells *in vitro* by mimicking two strong activation and survival signals: an activated B cell receptor (BCR), usually activated by the antigen, and a CD40 mediated one, respectively, usually mediated by a helper T cells (fig 1.11). The immortalized cells are often referred to lymphoblastoid cell lines or LCL. At this point the viral genome is not linear but is maintained as a circular episome, being distributed equally to daughter cells at each cell division [161]. Interestingly, this oncogenic virus has the ability to be found also as integrated copies in the host's chromosomes. This is particular evident for a Burkitt's Lymphoma cell line, named Namalwa, which mainly contain only integrated copies of EBV [178]. The study of integration copies is of great interest because this event can directly contribute to oncogenic mutations, having multiple integration sites distributed mainly on repeated regions (G-Band-Positive) but also functional genes areas (cell motility factor MACF1; cell differentiation/tumour suppressor BACH2, myeloid and B cell protooncogene REL and

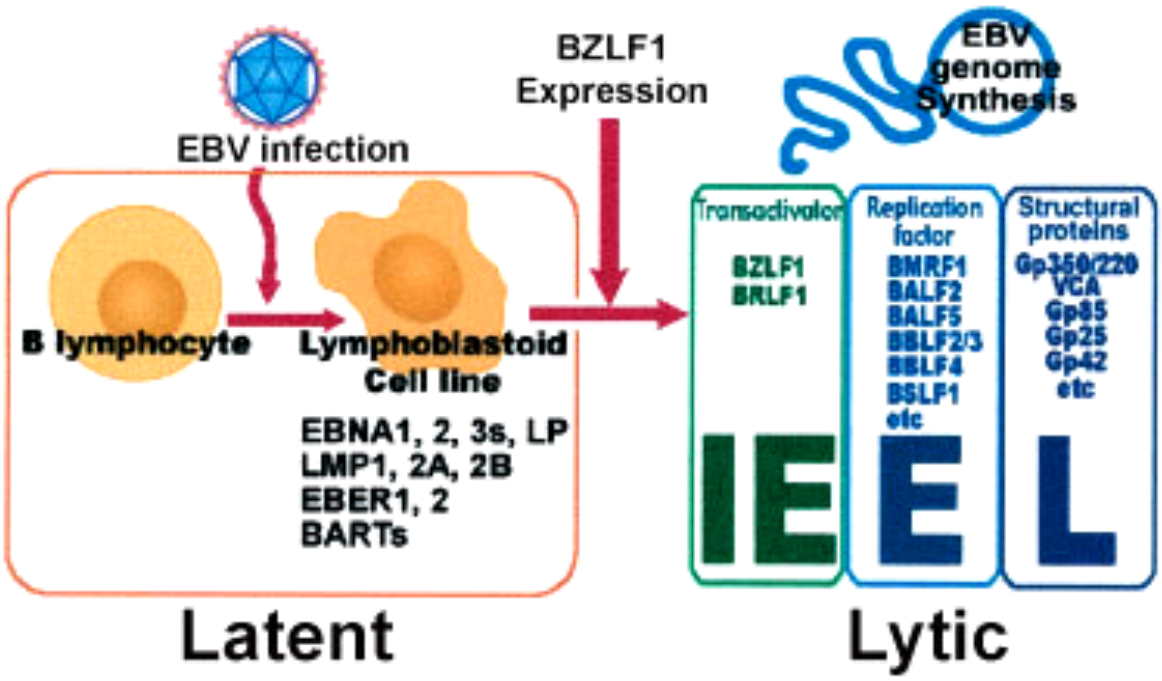


Figure 1.10. Epstein Barr Virus Life Cycle. EBV can infect Naïve B cell, pushing them into an iper-proliferative status (lymphoblast) which resemble the GC CD38+ stage. At this point it can survive in a silent form, pushing the cell toward a Memory B cell differentiation (named latency phase) with occasional cycles of replication (named lytic phase). During the latency phase EBV infected B cells are characterised by a distinct EBV gene expression profile including EBV nuclear antigen (EBNA), latent membrane protein (LMP) and Epstein–Barr virus-encoded small RNAs (EBER). Instead during the lytic phase the viral machinery switch to different molecules in order to reproduce the structural elements of the virus, such as BZLF or the Viral Capside Antigen (VCA). From *"Tsurumi, T., M. Fujita, and A. Kudoh, Latent and lytic Epstein-Barr virus replication strategies. Rev Med Virol, 2005. 15(1): p. 3-15"* [179].

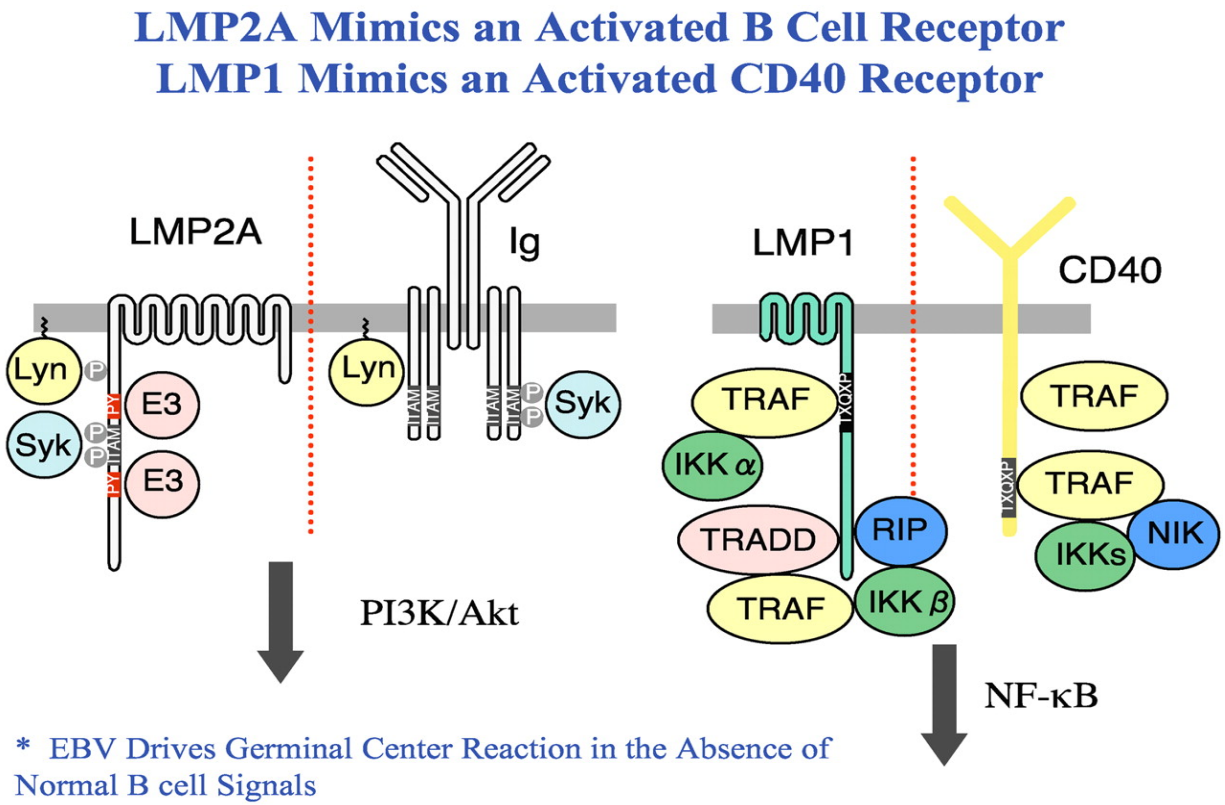


Figure 1.11. Epstein Barr Virus Immortalizing Machinery. LMP1 and LMP2A mimic CD40 and Ig receptors in B cells. LMP1 activation induce canonical and non canonical NF-κB signaling. On the other side LMP2A activates B cell by means of the Src family tyrosine kinase Lyn with a final results of a constant phosphorylation of LMP2A, which allows the recruitment of Syk (TNF receptor associated factors (TRAF); TNFR1-associated death domain (TRADD) receptor interacting protein (RIP)) From *"Gordon, L.I. and R. Longnecker, Off-targeting oft-targeted CD20 in cHL. Blood, 2012. 119(18): p. 4095-6".* [180].

b-cell lymphoma protein 11A (BCL11A) [181]. In order to distinguish between episomal and integrated copies, Reisinger et al have shown that fluorescence *in situ* hybridization (FISH) combined with dynamic molecular combing of the host cell DNA can facilitate this discrimination [182].

Thanks to LCL, it has been possible to investigate *in vitro* the different four latent programs, which also depend on the differentiation state of the infected cells: i) Latency III (or growth program), probably the most described one, where naïve B cells in tonsil are targeted, all latent genes are expressed driving a great proliferation of the B cells (EBER1-2, EBNA 1-LP, LMP1,2A,2B); ii) Latency II or default program, usually activated in the germinal center (EBER1-2, EBNA-1, LMP2A, LMP1); iii) Latency I or Latent program, characteristic of the memory cells in the blood stream (EBER 1-2 and EBNA-1); iv) Latency 0, EBER and LMP2A.

Basically, once the naïve B cells in the lymphoid tissue of the Waldeyer's ring become infected, the virus activates the cells to become proliferating blasts (Latency III). Then, switching between the different programmes, the cells will enter inside the follicle to initiate a GC reaction (Latency II). Here they will receive surrogate survival signals thanks to the LMP1/LMP2A axis, with the final aim of finally differentiate into a long-lived memory B cells (Latency I/0), in orders to survive protected into the blood stream. This is called "The germinal centre model" (GCM) and is supported by several evidences [183] (Fig 1.12). Quoting Thorley-Lawson from the Epstein Barr Volume 1 (Münz, 2015) *"the GCM provided a way to understand the complex biology of EBV. It has stood for 15 years and many tests of its reliability and predictive power. To date, it remains the only model that consistently provides a conceptual framework for understanding the complex and subtle behaviours of the virus".* It is important to mention that some authors have highlighted the possibility where also memory cells could become directly infected and that persistence could happen without transformation outside of the germinal center [184]. For instance, it has been shown by Babcock et al that the

frequency of infected memory B cells expressing the latency 0 in circulation is higher in IM than in persistent infection [185].

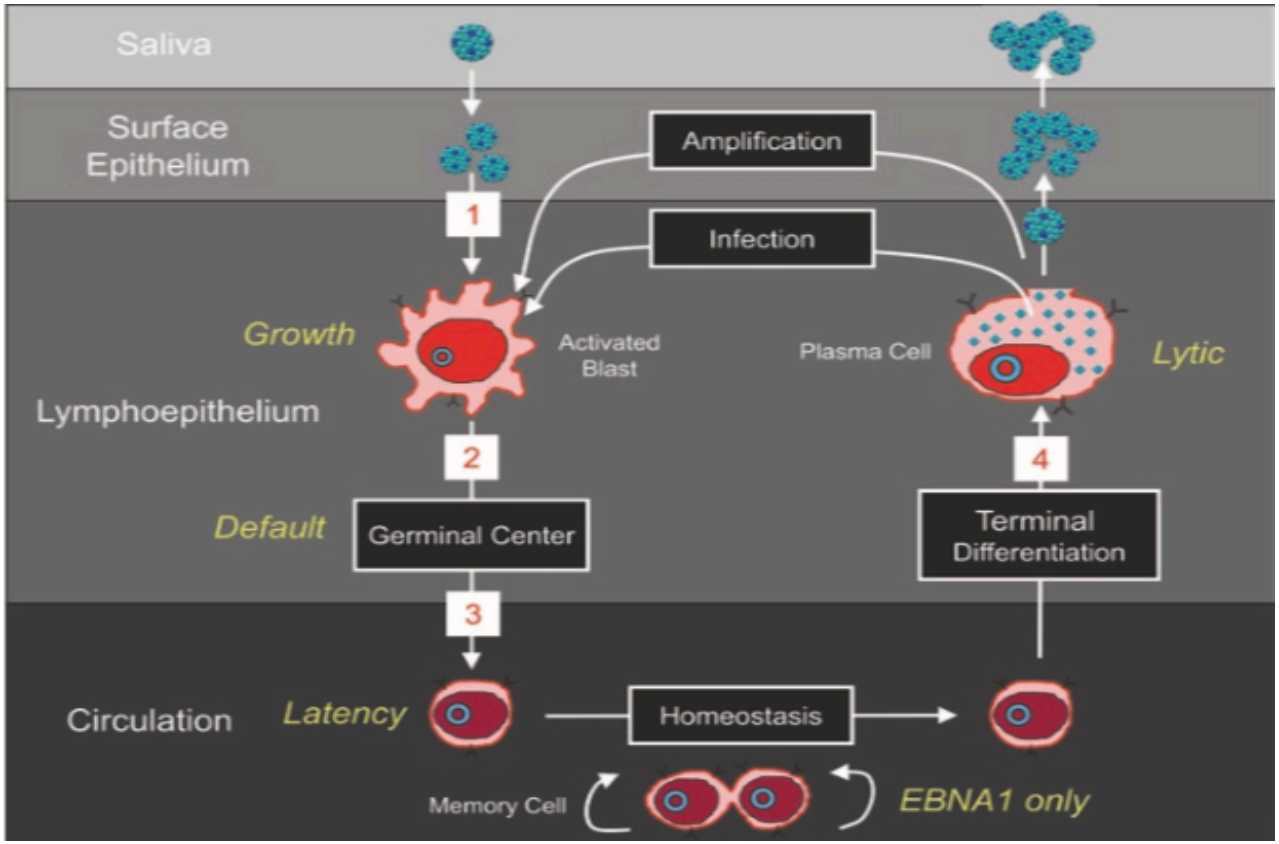


Figure 1.12. The germinal Centre Model: This Diagram has been obtained from "EBV persistence-Introducing the virus Thorley-Lawson (From *"Epstein Barr Volume I, Münz, 2015"*). For details check text.

EBV and latent antigens

Apart from the one already described, other tools essential for the molecular viral machinery are EBNA-1, EBNA-3C and EBNA-2 [186]. In order to become effective all these elements are sequentially orchestrated in their expression by different promoters. EBNA-1 binds the EBV genome to chromosome mediating episomal persistence and maintenance. EBNA-2 is a transactivator of both viral genes (LMP1 and LMP2A) but also cellular genes (such as CD21, c-FGR, c-MYC and specifically CD23). Indeed one of the basic feature of EBV is the presence of the 3-Kb BAMHI W repeat region (fig. 1.13), a major internal repeat containing tandemly arranged copies of similar size sequence and exon content [187]. Essential for the transforming function, each BAMHI repeat contains Wp, the first promoter to be activated upon B cells infection, and W1 and W2, two exons which encode

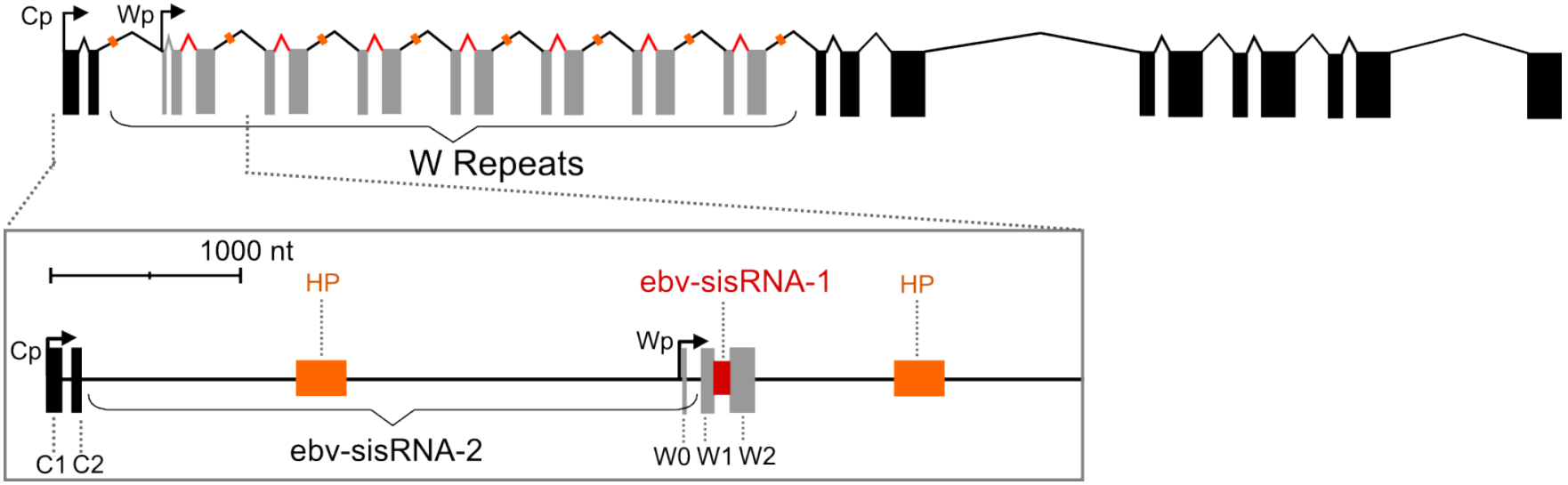


Figure 1.13. BamHI W Repeated region. EBV latency III transcript showing location of EBV W repeat stable introns. From: *"Moss, W.N., Analyses of non-coding RNAs generated from the Epstein-Barr virus W repeat region. Proceedings Iwbbio 2014: International Work-Conference on Bioinformatics and Biomedical Engineering, Vols 1 and 2, 2014: p. 238-252"* [188] .

for a 66-amino acid domain for EBNA-LP, the first EBV nuclear antigen expressed [189]. Together with EBNA-2, these antigens trigger the full latent gene expression activating Cp, another promoter whit a larger activation activity (which block Wp and produce EBNA1,2,3A,3B,3C) (fig. 1.7). At this point also the LMP promoters are activated, thus producing the mRNA for LMP1 and 2 [190]. It is important to add that the cytoplasmic portion of LMP1 contains carboxyl terminal activator regions (CTARs) that are similar to the human CD40, which are able to interact with the adaptor proteins tumour necrosis factor (TRAFF) and TNF receptor-1 associated death domain (TRADD) [191]. On the other hand, LMP2A has a N-terminal cytoplasmic portion with eight tyrosine residues, two of which form an immunoreceptor tyrosine-based activation motif (ITAM), the same as the one found in the α and β chains of the B cell receptor [192].

EBER1 and 2, first discovered in 1981, are small non-polyadenylated RNAs that represent the most abundant RNA in cells during latency. They are short nucleotides in length, respectively 167 and 172 nt. Intriguingly, they are transcribed by the host's RNA polymerase III during the latent infection. Many of the function of EBERs are still unclear but it is known that they can bind the ds-RNA-activate protein Kinase (PKR), which works in synergy with interferons, suggesting a mechanism in which the EBERs can enhance viral resistance [193]. The expression of EBER can induce cancer in severe combined immunodeficient mice (SCID) [194, 195]. Specifically, EBER1 can associates with human ribosomal protein L22, causing the latter to move from the nucleolus to the nucleoplasm [196]. Also, it has been shown by Kitagawa that this molecule can induce IL-10 production in Burkitt's Lymphoma [197].

EBV and lytic antigens (reactivation):

The complete infectious viruses are produced during the lytic phase, as usually been observed during acute infection and sporadically in persistent infection. As already mentioned, the process is kicked into gear by BZLF1 and BRLF1, resulting in the expression of essential early DNA synthesis molecules (DNA polymerase) such as BMRF1, BHRF1, BFRF1 and BALF2. Different phenotypes and genetic polymorphism of these genes from isolated EBV specimen account for the enhanced lytic replication capacity that has been shown for the M81 isolated cells in a NPC patient, providing evidence that - despite EBV almost global infection - small differences in strains can strongly contribute to oncogenesis [198].

Late lytic genes are then expressed in order to produce the protein need for viral structure, like BLLF1and BALF4 (who encodes respectively for the envelope glycoproteins gp350, gp220, and gp110).

At this point the real replication - a series of events that will also lead to the host cell death – will starts, with the synthesis of the viral DNA inside the nucleus. The DNA is then packaged insides the viral capsids and released into the cytoplasm, ready to fuse with the host plasma membrane in order to be spread from the host cells in an enveloped form.

A complete list for both latent and lytic genes and proteins can be found respectively in table 1.7 and 1.8.

Latency-associated genes	Proposed function
Epstein-Barr virus nuclear antigen 1 (EBNA-1)	Binds EBV genome to chromosomes mediating episomal persistence and maintenance; may inhibit apoptosis in Burkitt lymphoma cells; Gly-Ala repeat inhibits processing of EBNA1 through protesomes, precluding presentation of antigenic peptides with MHC-I at the cell surface, thus avoiding detection by CD8+ T cells
Epstein-Barr virus nuclear antigen 2 (EBNA-2)	Transactivator of several viral genes (LMP1 and LMP2A) and cellular genes (CD21, c-FGR, c-MYC, CD23) essential for cellular transformation ; interacts with transcription factor RBP-JK to promote conversion of resting B lymphocytes into lymphoblastoid B cells.
Epstein-Barr virus nuclear antigen 3C (EBNA-3C)	Essential for EBV-mediated transformation of primary B lymphocytes and interacts with RBP-JK ; EBNA-3C promotes LMP1 expression in the presence of EBNA-2.
Epstein-Barr virus nuclear antigen leader protein (EBNA-LP)	Key role in upregulating cell gene expression critical for lymphoblastoid B cell outgrowth and augments the ability of EBNA-2 to transactivate LMP1.
Latent membrane protein 1 (LMP1)	Mimics signaling through CD40, a B cell activation and differentiation receptor, and in turn activates NFκB pathways to promote cell survival; promotes B cell lymphoma formation in transgenic mice expressing the gene.
Latent membrane protein 2A (LMP2A)	Constitutive dominant negative modulator of B cell receptor signaling by sequestering tyrosine kinases through interactions with its immunoreceptor tyrosine-based activation motifs, preventing entry into the lytic cycle by antigenic stimulation; promotes B cell survival through activation of the same sequestered tyrosine kinases in an antigen- and B cell receptor-independent manner.
Epstein-Barr virus–encoded RNAs 1 and 2 (EBER1 and 2)	Small nonpolyadenylated RNAs that represent the most abundant RNAs in latently infected cells; role in EBV infection still unclear

Table 1.7. EBV virus-encoded latent infection genes products. Readapted from *"Spectrum of Epstein-Barr Virus-Associated Disease, J.L. Kutok and F.Wang, 2006, Annual Review of Pathology: Mechanism of Disease"* [161]

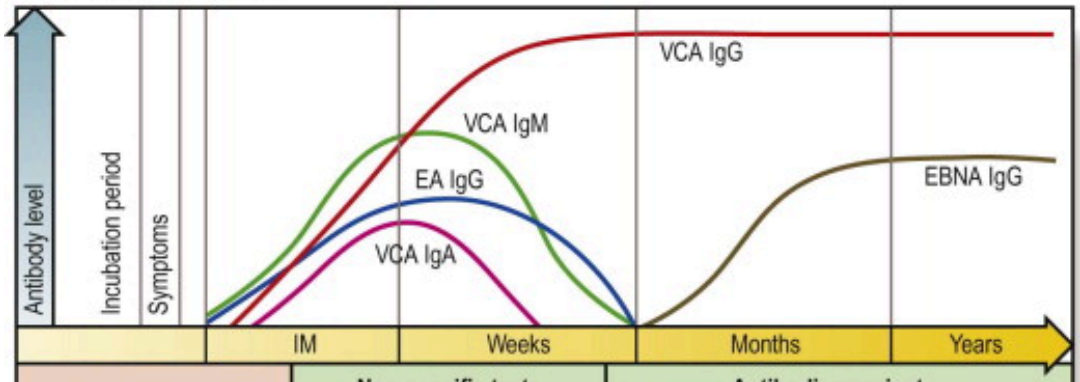
Latency-associated genes	Proposed function
BZLF1 / BRLF1	Key transactivators of lytic EBV gene expression, coordinately upregulating expression from early EBV promoters
BHRF1	Colinear homology with BCL-2, prevents apoptotic cell death in lytic EBV infection; nonessential for growth transformation of B cells and for virus replication in culture
BCRF1	Homology to human interleukin-10; may downregulate the host immune response during viral replication
BALF5	Replication factor; core DNA polymerase
BALF2	Replication factor; single-strand DNA binding protein
BMRF1	Replication factor; processivity factors
BSLF1 and BBLF4	Replication factor; primase and helicase complex
BBLF2/3	Replication factor; spliced primase helicase complex component
BKRF3	Replication factor uracil DNA glycosylase
BLLF1 (gp350/220)	Mediates adsorption between EBV and CD21

Table 1.8. EBV virus-encoded lytic infection genes products. Readapted from *"Spectrum of Epstein-Barr Virus-Associated Disease, J.L. Kutok and F.Wang, 2006, Annual Review of Pathology: Mechanism of Disease"* [161].

Control of EBV infection by the immune system

The sporadic expansion of the virus – as shown in the GCM – is counterbalanced by the immune system where both humoral and cell mediated adaptive responses take place. Indeed, several antibodies against different viral elements can be found in the serum of EBV+ individuals, such as those against EBNA, VCA (viral capsid antigen), EA (early antigen). These antibodies are not just helping in containing the spreading of infection but are also a consistent and reliable diagnostic tool to discriminate between different EBV-related diseases or phase of infection (see table 1.9) [199].

Regarding the cell mediated response the main guardians of EBV homeostasis are the cytotoxic CD8+ T cells, which are able recognise EBV infected cells upon lytic reactivation but can also contribute actively to patients' symptoms. Things are more complicated during the various latency programs: the CD8+ specific T cells for latent antigen are present but to a smaller degree. The majority of them are targeting EBNA-3A, 3B and 3C [200]. Also, subsets targeting LMP2A and LMP1 are present [201]. It is important to report that a response to EBNA-1 protein has been described in IM patient carrying particular HLA alleles, specifically HLA-B*3501, which are usually uncommon in the general population [202]. People with these alleles are characterised by a strong CD8+ T cell response against EBNA-1. However, CD8+ are mostly unable to recognise EBV infected cells during latency I (characterised by EBNA-1 expression). This account for the virus persistence a low level in infected individual (between 1-5 cells every 100'000-1'000'000 cells) [163]. Also NK cells can be involved in the EBV response through the type I INF axis, and this is quite evident in NK deficiencies, which are usually associated with increasing susceptibility to EBV infection [203].



Clinical entity	Non-specific tests		Antibodies against					
	Heterophylic antibodies	Atypical lymphocytes	VCA			EA		EBNA1
			IgM	IgG	IgA	EA(D)	EA(R)	IgG
No infection	–	–	–	–	–	–	–	–
Infectious mononucleosis	+++	+++	++	+++	+	+	(±)	–
Past infection	–	–	–	++	–	–	(±)	++
Reactivation	–	–	(+)	+++	++	++	+	(+)
Burkitt lymphoma	–	–	–	+++	–	(±)	++	+
Nasopharyngeal Ca	–	–	–	+++	++	+	(±)	++

Table 1.9. EBV Diagnostic Test Profile. From *"Virology, Epstein-Barr Virus, Stephen N.J. Korsman, 2012"*.

Subpopulation of CD23/CD58 B cells that proliferates *in vitro* after infections

It is quite clear at this point that, by these molecular means, B cells acquire a lymphoblastoid cell lines phenotype, which resemble a germinal center B cell for the expression of CD38, CD10, CD77, BCL-6 and AID, plus the chemokine receptors CXCR4, CXCR5 (but not CCR7) [204]. These cells are actually capable of indefinite growth. Moreover, *in vitro* characterisation by Megyola et al of EBV+ B cells from healthy donors have shown that, after infection, these lymphoblast are mainly formed by two specific and unusual subsets: a $CD58^{+}/CD23^{low}$ non proliferating IL-6 producing subset and a $CD58^{+}/CD23^{hi}$ proliferating one [205]. CD23, also known as Fragment crystallization (Fc) epsilon RII (FcεRII) is the "low-affinity" receptor for IgE while CD58 or lymphocyte function-associated antigen 3 (LFA-3) is a cell adhesion molecule expressed on APC, particularly macrophages. There has not been any functional connection between EBV and the purpose for these receptors, nevertheless CD23 seems to be highly expressed in EBV-related lymphomas.

EBV in Autoimmunity: Focus on MS

Due to the ability to infect preferentially B lymphocytes - one of the most important actors of the immune system – the EBV has nourished a wide field of speculation regarding a possible connection to autoimmunity. Several studies have been done in this direction but a clear a demonstration of causality, defined as state in which the reduction of the frequency of the virus results in a reduction of a specific autoimmune disease or condition, has bene difficult. In fact, this would require that prevention of the infection results in a reduction of the disease. As suggested by Ascherio, a strategy to understand this could be trough i) meta-analyses combining the data from several case–control studies ii) case-control studies of paediatric populations in which prevalence of EBV positivity is still low at ~75 % or

less. The two most studied diseases are SLE and MS and, so far, evidence has been found only for the latter. For this reason, the next chapter will focus only on MS.

MS is a chronic inflammatory and neurodegenerative diseases that affect mainly the central nervous system and is caused by an immune-mediated demyelinating process. Several immune cells (B, T, macrophages) are usually found in biopsies material [206]. Alteration at a systemic level have being described, such an altered cellular immunity against myelin antigens or impaired response of Treg [207-209]. The disease usually has a higher incidence in women more than man and manifest between 20 and 40 years of age. Geographically, the most affected areas are Europe, the USA, Canada and Australia. A combination of genetic factors (the risk-allele HLA-DR15) and environmental factors (vitamin D deficiency, cigarette smoking, IM and EBV infection) has been highlighted in MS [210, 211].

Regarding the epidemiological evidence, we now know that MS risk is 2.3-fold-higher between individual with a clinical history of IM, as compared to individual which have never been affected by the disease. It has been speculated that high level of hygiene in childhood can reduce exposure to environmental agents and this can delay the age of primary EBV infection. Furthermore, individual which are EBV negative tend to not develop MS, although they have high seroconversion rates and high IM incidence when exposed to the virus [212].

Regarding a potential mechanism relating EBV to MS, it has to be mentioned that: i) EBV viral load in patient with MS is only modestly increased; ii) MS has not been reported as a complication of immunosuppression in post-transplant lymphoproliferative diseases; iii) monoclonal antibodies that either prevent T central migration into the CNS (natalizumab) or deplete B cells (rituximab) are both very effective against MS relapses [213]; iv) anti-EBNA-1 and EBNA-2 have a strong correlation with MS development in young adults [214]. These findings suggest that a hyperactive or uncontrolled immune response toward EBV operated by the immune system - rather than lytic viral replication or over proliferation of

infected cells - can cause MS. Several epidemiological and clinical evidence have shown that anti-EBNA-1 antibodies and $CD4^+$ EBNA-1 specific T cells can cross-react with myelin antigens and leading to the evident damage present in MS patients [215]. Also, Serafini and Aloisi just recently demonstrated that EBV specific CD8+ T cells can access the brain and selectively target infected cells, strongly contributing to MS pathogenesis [216]. It is also believed that B cells can take part in this process not only in antibody production but also with the presentation of cross-reactive antigens [217].

Rheumatoid Arthritis and the role of EBV: evidences and controversies

Overview:

The rationale for the association between EBV in RA (and in general in other autoimmune disease) is that viral latent genes *in vivo* may potentially rescue an autoreactive B cell clone - which in a normal scenario would be induced either to apoptosis or to clonal anergy during a series of physiological tolerance checkpoints - allowing it to access the GC and, subsequently, enter the blood stream as part of the memory B cell compartment. This latently infected B cell is not able to produce antibodies by itself because, upon the intrinsic plasmacells final differentiation, it would start to produce viral particles and become a target for T CD8+ anti-EBV specific cytotoxic cells. Nevertheless, an EBV-rescued B cell could theoretically escape depletion and present autoantigens to T cells favouring a breach of immunological tolerance. Interestingly, some evidence supporting this theory has been provided by Swanson-Mungerson and Longnecker in a mouse model of autoreactivity, where they have demonstrated that LMP2 can actually breach tolerance [123].

Unfortunately, epidemiological data linking EBV and RA are debatable since EBV viral infection is by definition not more frequents in patients

with RA than in control patient [218] given that 95% of the population is infected. Experimentally, one of the most important studies about this topic has been done produced by Tracy et al, where they profiled the self and polyreactive properties of BCRs expressed by EBV infected memory B cells present during the acute phase of infection with the final aim of understanding the role of the virus in this process. In their elegant work, they have discovered that EBV can be found in self-reactive memory cells during IM, but it does not preferentially survive in these cells. Of more importance, apparently EBV is able to persist in B cells producing low affinity self-reactive and polyreactive antibodies, and this might be connected to a mechanism in which the virus ensure its survival by reducing the chances of being seen as a threat by the immune system [219].

However, in my laboratory Croia and Bombardieri have generated an impressive amount of data connecting EBV with RA and SS and linking it to a possible autoimmune mechanism in ectopic GCs. Also new evidences are emerging every year about this issue or, in general , about the role of EBV other autoimmune disease, such as the above mentioned work published by Serafini and Aloisi in multiple sclerosis in 2019 [216]. It is important to mention that one of the most comprehensive review has been published by Balandraud and Roudier in 2017, and some of the most important point highlighted in it will be discussed here [64].

Molecular mimicry mechanism between EBV and self-proteins

As already stated in this introduction, there are several points of connection between EBV epitopes and autoantigens. This molecular mimicry could potentially break tolerance and lead to the development of RA. The most likely candidate is the shared epitope, since part of the QKRAA sequence is present in the gp110, a glycoprotein expressed during replication of the virus on the envelope [220]. Indeed, EBV seropositive healthy individual

have T cells who recognise QKRAA and their antibodies are able to bind gp110.

Regarding the HLAs complex, the similarities are not limited to HLA-DRB1 since Fox et al in have later showed that in SS, the protein expressed by the allele HLA-DQ*0302 can actually overlap EBNA-6 aminoacidic sequence. Also, authors have showed that deiminated EBNA-1 can be a target of ACPA and can cross react with form of human citrullinated fibrin [221, 222]. It is interesting to mention that recently Harley et have shown by computational method that several autoimmune diseases risk loci are occupied by EBNA-2 [223].

Humoral response to EBV antigens in rheumatoid arthritis

It is fascinating that one of the most important antibodies against EBV, the anti-EBNA-1, was discovered in RA an initially named rheumatoid-arthritis-associated nuclear antigen (RANA) [224]. These antibodies were targeting nuclei antigens in B cells and, most importantly, RA patient showed a higher reactivity (67%) compared to healthy control (8%)[225, 226]. Was later discovered that this mysterious antigen was, in fact, EBNA-1 [227]. Another strong connection was made by Alspaugh, who demonstrated that not just anti EBNA-1, but also other antibodies targeting VCA and EA, were significantly increased in serum and synovial fluid from RA patient[225].

It has been already mentioned that viral antigens can be citrullinated and become an ACPA target and Merlini et al identified a small region of 23 aa (between aa 35 and 58) on the EBNA-1 that is highly reactive with serum from RA patient [228]. Pratesi and Migliorini later confirmed - using synthesized peptides modified or unmodified for citrullination - that this region was indeed target by ACPA [221]. Furthermore it has been demonstrated that citrullinated fibrinogen peptides can be targeted by anti EBNA-1, casting further evidence on how the virus can break tolerance trough antigens-cross reactivity [222]. More recently Trier et al have

showed that citrullinated EBNA-2 peptides from three different strains of EBV (B95-8, AG786 and GD1) are targeted very differently by RA sera. For each strain several peptides have been produced and then tested with sera from RA patients and healthy donors (Fig.1.14). Specifically, one strain, the AG786, react more than the others, pointing at the hypothesis that RA may be caused by a specific strain of EBV.

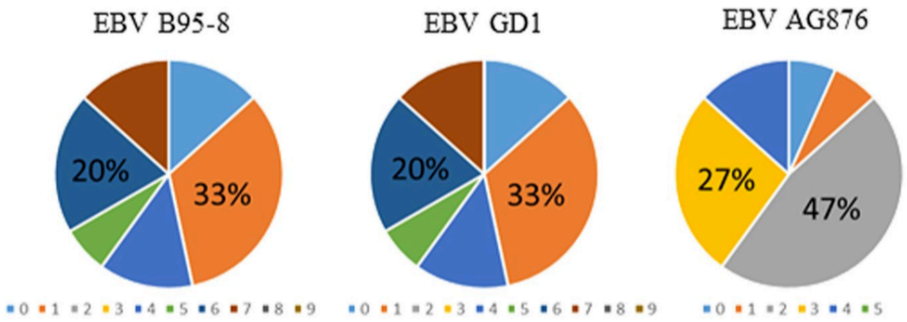


Figure 1.14. Overlapping antibody reactivities to Cit-Gly-containing peptides among the three Epstein-Barr virus strains B95-8, GD1 and AG876. From "*Antibodies to a strain-specific citrullinated Epstein-Barr virus peptide diagnoses rheumatoid arthritis, Trier et al, 2018, Scientific Reports: "The coloured fields represent the number of peptides recognised for the individual sera, and hence the degree of overlapping reactivities in a serum pool"[229].*

Cell-mediated immune response to EBV in rheumatoid arthritis

Almost 40 years ago Tosato studies demonstrated that T cells from RA patients are defective in controlling the virus as they do in healthy donor. Indeed they showed that LCLs - the cells line naturally harbouring the virus - arise more often from RA blood than from controls, underlying an impairment from the CD8+ T cells in taking care of the infection [230].
Of interest, RA patients have less circulating anti-gp110 T cells, and this lower frequency has been connected to HLA-DRB1*04:04, one of the shared epitopes alleles, which has shown less efficiency in controlling the gp110 epitope [231]. Toussirot et al have later showed that this impairment it is also connected to an increased disease activity [232]. Other studies have highlighted in RA the clonal expansion of dysfunctional T CD8+ anti-EBV in responding to lytic EBV peptides [233].

EBV load in patients with RA:

RA patients are characterised by high systemic EBV load, directly connected to the above-mentioned T cells impairment. Also, EBV DNA can be detected more frequently in PBMCs saliva and synovial fluid from RA. Several techniques have been used during the last 40 years in order to quantify EBV DNA (ISH, PCR, ecce cc) but the actual state of the art is real time PCR for BamHI W repeated region. Using this technique Balandraud and his group have managed to demonstrate that EBV load was 10-fold higher in RA patient than in healthy controls. Specifically they observed that the EBV DNA load was significantly increase PBMC from RA patients(n = 84, mean 15.6 copies/500 ng DNA) than in non-RA inflammatory diseases (n = 34, mean 5.7 copies/500ng DNA) and normal controls (n=69, mean 1.9 copies/500 ng DNA)[234].
RA patients usually undergo long-term treatments, and some of these drugs have shown to have an effect on EBV load. Some speculations have been made that long-term use of methotrexate might lead to induction EBV

associated lymphomas, but further studies are needed[235]. The majority of TNF blockers – such as Infliximab or Etanercepet – shows that patient load is stable over time[236]. Nevertheless, new biological treatment such as abatacept (CTLA4 Ig) and the anti-IL-6 Tocilizumab are under study. Regarding abatacept, the fact that this drug inhibit T cells activation might lead to increase EBV replication, as found by Bassil in patients undergoing kidney transplant [237].

On the other hand, Erre et al found by PCR that EBV load was increases in RA patient compared to control, but surprisingly lower in RA patient undergoing treatment with tocilizumab[238]. This might be connected to the evidences highlighted by Megyola that different subsets of EBV infected cells might sustain each other through an IL-6 mechanism, leading to higher levels of circulating infected B cells.

EBV in the ectopic GC of synovial membrane of RA patients

Regarding the possibility that EBV might survive in ectopic GC inside the synovia of RA patient, several evidences have been provided by Croia et al in my lab. The authors investigated – through RT-PCR, immunohistochemistry and in situ hybridisation - latent and lytic infection in 43 synovial tissue characterised by presence or absence of ectopic GCs and compared them to 11 control osteoarthritis. This broad screening has analysed LMP2A and EBER transcript, EBER+ cells, and immunoreactivity for EBV latent (LMP1 and LMP2A) and lytic (BFRF1) antigens. Intriguingly, they highlighted a strong EBV dysregulation in RA patient positive for ectopic GC, which was not identified in osteoarthritis controls. The most important result is that EBV infection was colocalizing in synovia with autoreactive ACPA+ B cells and plasmacells (Fig.1.15). Further *in vivo* evidences have been provided by transplantation of RA synovial tissue positive for ectopic GC) into SCID mice, which resulted in a strong production of ACPA and anti-viral citrullinated peptides 1 and 2 (VCP1 &VCP2). Also, analysis of T cells (both CD4+ and CD8+)

suggested that EBV persistence in the ectopic GC may be possible by the exclusion of T cells from the B cells follicles. A similar work has been repeated by Croia et al in SS, confirming that EBV dysregulation was present only in ectopic GC positive salivary gland and was colocalizing with plasmacells displaying Ro 52 [239].

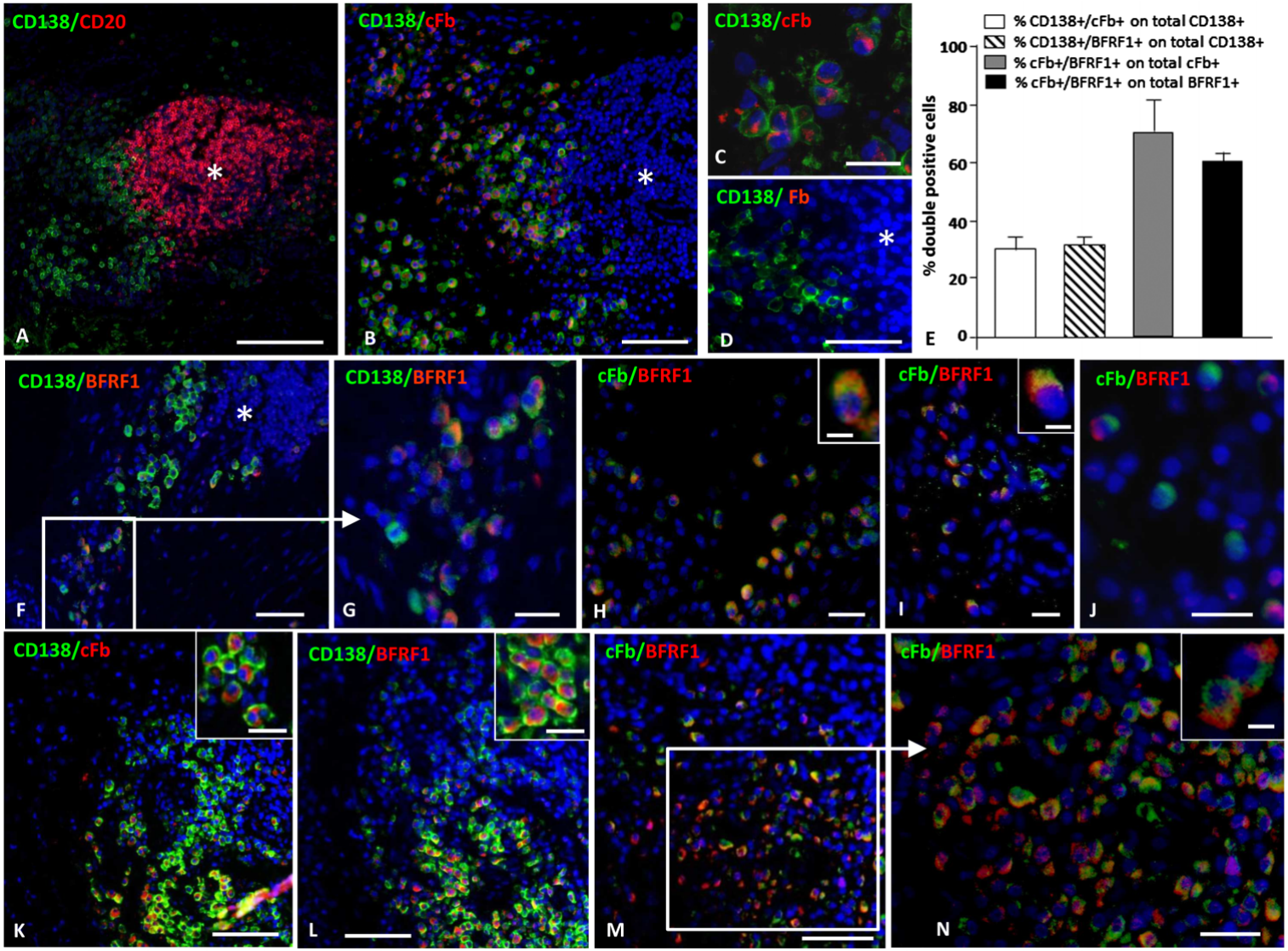

CD138/CD20
CD138/cFb
CD138/cFb
CD138/ Fb
% CD138+/cFb+ on total CD138+
% CD138+/BFRF1+ on total CD138+
% cFb+/BFRF1+ on total cFb+
% cFb+/BFRF1+ on total BFRF1+
% double positive cells
100
80
60
40
20
0
A
B
C
D
E
CD138/BFRF1
CD138/BFRF1
cFb/BFRF1
cFb/BFRF1
cFb/BFRF1
F
G
H
I
J
CD138/cFb
CD138/BFRF1
cFb/BFRF1
cFb/BFRF1
K
L
M
N

Figure 1.15. Detection of the Epstein–Barr virus early lytic protein BFRF1 in anticitrullinated protein/peptide antibodies-producing plasma cells in ectopic lymphoid structures+ rheumatoid arthritis (RA) synovia (previous page): From Epstein–Barr virus persistence and infection of autoreactive plasma cells in synovial lymphoid structures in rheumatoid arthritis, Croia et al, 2013, Basic and Translational Research. "*Double immunofluorescence staining for CD20 (red) and CD138 (green) shows a representative B-cell follicle surrounded by plasma cells (A). At the border of the same B-cell follicle (marked with an asterisk in A,B,D,F) many of the CD138 plasma cells (green in B–D) react with citrullinated fibrinogen (cFb; red in B,C), indicating that they produce APCA. No binding of non-citrullinated Fb is detected (D). In the same perifollicular area, a substantial fraction of CD138+ plasma cells (green) coexpresses BFRF1 (red) (F) the area marked with a frame in (F) is shown at higher magnification in (G) while most BFRF1+ cells (red) are cFb-reactive (green), and vice versa (H–J). The insets in (H,I) highlight the colocalisation of cFb and BFRF1 in the cell cytoplasm. Panels (K–N) show double immunostainings for CD138/cFb (K), CD138/BFRF1 (L), and cFb/BFRF1 (M,N) the area marked with a frame in (M) is shown at higher magnification in (N) in consecutive sections of a different RA synovial sample. The column bar graph in (E) shows the results of cell counts performed in six synovial RA samples. Values are means±SEM of the percentages of double positive CD138+/cFb+ and CD138+/BFRF1+ cells on total CD138+ plasma cells, and of double positive cFb+/BFRF1+ cells on total BFRF1+ and cFb+ cells. Bars: 100 mm in A; 50 mm in (B,D,F,K–M); 20 mm in (C,G,H–J,N) 10 mm in the insets in (K,L) 5 mm in the insets in (H,I,N).*"

CHAPTER 2 RATIONALE OF THE THESIS AND AIMS

The main aim of my PhD project is to clarify the intimate relationship between EBV, autoreactive B cells activation and the maintenance of autoimmunity against citrullinated antigens in the joints. Given the strong connection between EBV and RA I hypothesized that recreating in vitro the synovial environment using a RA-FLS feeding layer could enhance the proliferation of B cells, due to the production of survival factors. I also hypothesized that EBV-infected B cells might have a proliferative advantage compared to their uninfected counterpart and that this might lead more easily to the escape of an autoreactive B cells, which could resemble a possible mechanism that triggers autoimmunity in vivo. In this sense the RA-FLS could create a unique niche for EBV-infected cells, increasing the chances of developing autoreactivity, due to lack of control.

In particular, I shall study if and how synovial fluid and synovial tissue RA-FLS enhance selectively the long-term survival of naturally occurring EBV infected B cells obtained from circulating B cells of RA patients and healthy donor. Also, I shall investigate if in turn EBV B cells are able to enhance RA-FLS inflammatory response. Finally understand if these cells are autoreactive.

Aim & Objectives

My PhD project aims to establish the relevance of EBV in sustaining autoreactive B cells activation in the RA joints. In order to do this, I need to:

i) Verify whether B cells forming long-term clusters in co-culture with RA-FLS display a preferential selection of clones harbouring EBV.

ii) Verify whether B cells harbouring EBV have an autoreactive phenotype and produce autoantibodies against RA-associated self-antigens.

CHAPTER 3 RESEARCH METHODOLOGIES.

Patient recruitment, sample process procedure and cell culture

FLS have been obtained from different sources: joint replacement surgeries, synovioctemy surgeries, synovial biopsies or synovial fluid of psoriatic arthritis (PsA), osteoarthritis (OA) and ACPA+ RA patients. B cell have been obtained from peripheral blood of the ACPA+ RA patients and healthy donor. All RA patients in the study fulfilled the 2015 ACR/EULAR classification criteria for RA. Patients enrolled in this study are part of the Experimental Medicine and Rheumatology Research Tissue bank (EMR-RTB, previously Qmul-Rbb). Patients sample have been collected at the Rheumatology department of the Mile End Hospital of London, part of the Experimental Medicine and Rheumatology (EMR) department of the Queen Mary University of London (QMUL). Healthy donor blood has been collected at the EMR department of QMUL.
Prior to sample collection a written informed consent was obtained from both patients and healthy donors. All tissue was obtained immediately after surgeries or collection. Table 4.1 is showing all sample collected. Table 4.2 is showing samples according to the relative experiment. Sample have not been tested for EBV seropositivity.

a) Isolation and Culture of FLS

Synovial fluid was collected in heparinised syringe, diluted upon arrival on phosphate buffered saline (PBS) and centrifuged 10' at 1500 rpm RT. The resulting pellet was resuspended in 10 ml of filtered complete growth medium (CM) [Dulbecco's Modified Eagle Medium: Nutrient Mixture F-12 (DMEM/F12) with 10% heat inactivated fetal bovine serum (FBS), 1% Antibiotic-Antimycotic (Streptomycin, Amphotericin B, Penicillin) (Gibco/Thermo Fisher Scientific)] and plated in a T25 tissue culture flask. Cultures were incubated at 37°C with 5% CO2 for 48 h, after which medium was aspirated and cultures were washed with (PBS) to remove

unattached cells. Every 10 to 14 days adherent confluents cells were removed from flasks by trypsinization, and splitted at 1:3 ratio in fresh growth medium. Complete growth medium was replaced every 3 to 4 days. After passage 4 cells were used for experiments in order to avoid any contamination from synovial macrophages, which are known to resist in culture during the first passages. After passage 4 FLS are the dominant cells resulting in a relatively homogenous population [240].

Synovial tissue was minced, washed with DMEM - F12 and plated in a 10 mm petri dishes with complete fresh medium. After 48 hours tissue was removed, mashed under a 70 μm nylon mesh cell strainer (Becton Dickinson, Oxford, UK) and digested overnight at 37°C with 1.5 mg/ml Dispase II in DMEM supplemented with 10% FCS, 50 IU/ml penicillin-streptomycin and 10 mM HEPES buffer (Gibco/Invitrogen, Paisley, UK). Suspension was then passed through cell strainer, centrifuged and the cell pellet resuspended in 10 ml of complete growth medium and plated in a T25 flask. Culture were maintained at 37°C with 5% CO2 and splitted at 1:3 ratio when they reached confluence. Passages 4 and 5 were used for experiments.

To have an estimation of the number of cells in culture, all FLS have been cultured and managed using seeding density rules found in literature (1.0-1.4 x 104 cells/$cm^{2,}$ Table 3.1). Seeding density appears to work better than standard counting cells protocol (like trypan blue) since the FLS – after being detached with trypsin - tend to form aggregates making it hard to have a reproducible count.

b) Peripheral blood mononuclear cells isolation and CD19+ purification.

Peripheral blood mononuclear cells (PBMC) were obtained from 20 ml blood sample collected from patients and healthy donors in heparinised tubes. PBMC were isolated using Hystopaque 1077 density centrifugation. CD19+ B cells have been purified from isolated PBMC by means of magnetic selection with CD19+ microbeads (Miltenyi Biotech, Germany) with a final yield between $1{,}8x10^5$ -$8x10^5$ number of purified cells. Cells were maintained in a suitable medium for B cell (CM-B: Roswell Park Memorial Institute-1640 (RPMI 1640) low glucose with 10% heat inactivated FBS, 1% L-Glutamine, 1% Antibiotic-Antimycotic). The purity of CD19+ B cells isolation was checked by fluorescence-activated cells sorting (FACS, BD Fortessa). Cells have also been characterized for immunoglobulin production. $2x10^5$ cells have been used for purity and characterization respectively.

Both fresh and frozen sample have been used in these experiments. It is important to state that for the co-culture system only frozen sample have been used. A comparison between fresh and frozen B cells samples to address differences in viability and proliferation has not been performed.

Co-culture Experiment:

In the first protocol FLS were plated in a 96 well plate (2'500 cells p/w) and cultured in FLS-CM for one/two weeks until sub-confluence (70-90%) was reached. For all these experiments FLS after passage 4 have been used. Before adding the PBMCs, FLS-CM has been completely aspirated. The ratio between cells has been modified based on the lowest number of mononuclear cells obtained during the purification (150'000 PBMC p/w).
Different stimuli have been added to the co-culture
i) IL-2 (1000 U/ml, Sigma) to enhance cellular survival.

ii) CpG (2.5 µg/ml, ODN2006, Invivogen) which is used to stimulate Toll Like receptor 9 in order to differentiate B cells toward antibody producing plasma-cells [241, 242].
Cells have been cultured in a total volume of 250 µl.

In the subsequent protocol optimization, I used a 48-well plate (7'500 cells p/w, 8 wells/condition) and cultured in FLS-CM for one/two weeks until sub-confluence (70-90%) was reached. For all this experiment FLS after passage 4 have been used.
Before adding the purified CD19+, FLS-CM has been completely aspirated. The ratio between cells has been modified based on the lowest number of mononuclear cells obtained during the purification (i.e., 7'500 FLS; 25'000 CD19+ p/w). 250'000 purified B cells have been stored again to be used as time 0 control in sorting analysis.
Same different stimuli have been added to the co-culture
i) IL-2 (1000 U/ml, Sigma) to enhance cellular survival;
ii) CpG (2.5 µg/ml, ODN2006, Invivogen)

At the beginning of the co-culture, cells and stimuli have been plated in a total volume of 700 µl (CM-B-cells). 150 µl have then been collected after two days and replaced with 650 µl in order to reach the final volume of 1200 µl. This difference at day 0 and day 2 was important to let the CD19+ cells seed to the bottom of the well and attach to the FLS during the first 48 hours (which are critical), but also to use a lower quantity of the two stimuli which, at the operational concentration, will not be consumed in the firsts days. 150 µl of SN has been then collected every 15 days (15, 30, 45, 60, 75, 90, 105 day) and replaced with 200 µl in order to compensate evaporation. SN were pooled between wells containing the same stimulus, centrifuged for 10' at 1800 rpm to obtain cell free suspension and frozen at -20C. As control, FLS alone (with or without stimuli) and CD19+ B cells alone have been used.

A maximum of 123 days has been reached but best conditions for this experiment has been assessed at 60-75 days. Matched sample (FLS and CD19+ cells from the same patient) was possible only in one case (co-culture of RA-FLS Qmul-Rbb 2018/003 and RACD19+ EMRRTB 0160).

In an optimization procedure for the co-culture system, different protocols and ratio between FLS and B cells have been tested in order to increase the likelihood of survival and growth of EBV-infected B cells. This has led to a new protocol of the co-culture in which 2 wells from a 12well plate are used as biological replicates with 30'000FLS and 360'000 B cells p/w respectively (2ml each of CM-B cells have been added). This increased the ratio between FLS and B cells from 1/1.6 to 1/12. Also, given the average distribution of infected EBV cell in normal population (1-5 cells every $1*10^6$ cells), the probability in the previous assay of obtaining an EBV infected cell was 0.50 while in the new one is 3.60. Furthermore, proliferation was observed earlier (14 days) and all experiment have been reduced to 28 days.

At least in one case (Experiment 0051), after sorting, EBV+ B cells have been plated again in a second co-culture and maintained for 36 more days, for a total time in co-culture of 64 days (fig 3.1). After this second co-culture, Experiment 0051 cells were clearly showing behaviour of LCL and from this point onward have been able to survive in culture without any feeding layer.

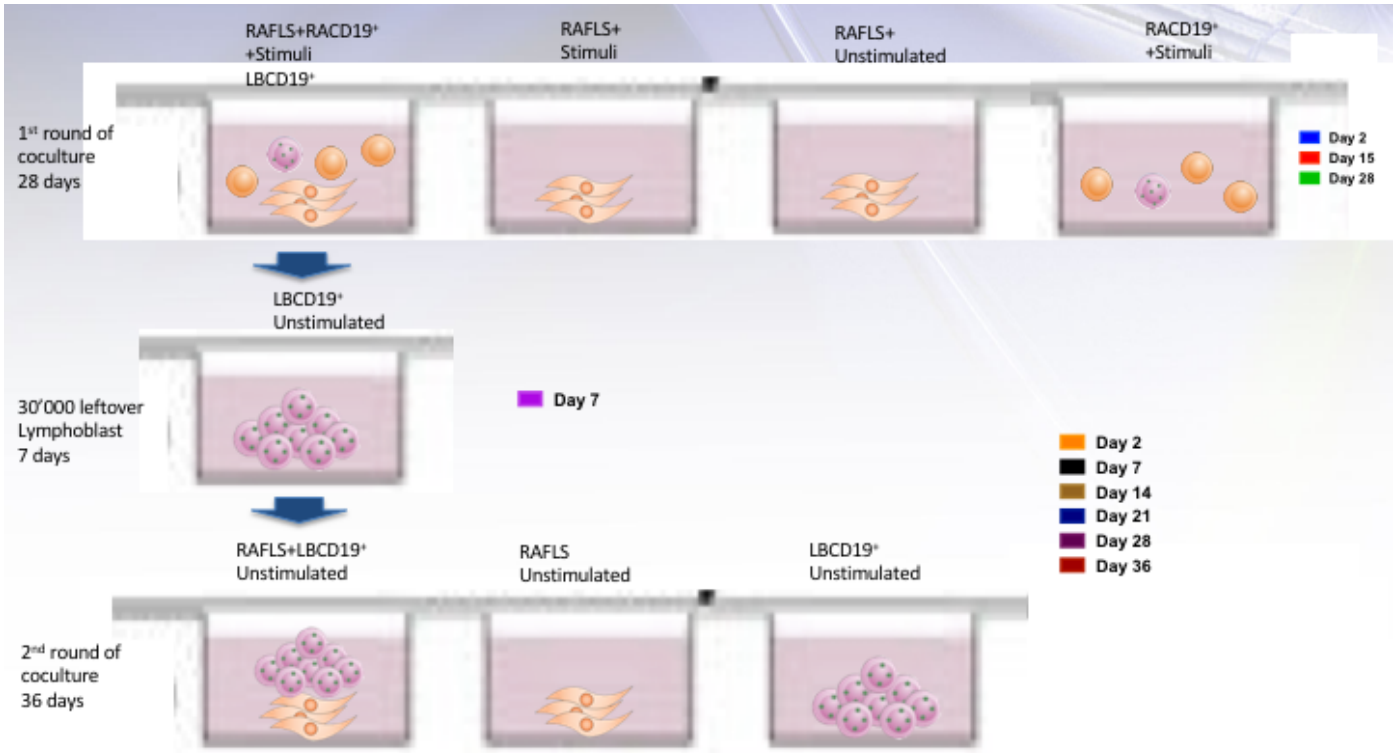


Figure 3.1. Experiment 0051: 1st and 2nd co-culture. Here is shown the protocol of experiment 0051. 4 condition from normal protocol are shown as previously described. After sorting not all of B cells have been used for BAMHIW qPCR. Instead 30'000 cells have been divided and plated again with or without fibroblast and kept in culture again for other 36 days for a total time in co-culture of 64 days.

Purity, Characterisation and Lymphoblast presence in Purified CD19+ B cells

All the panel have been designed using the Fluorofinder online platform, which allows to visualise the spectra in order to detect spectral overlap. An aliquot of the purified CD19+ B cells ($2x10^5$) have been screened for purity, characterized for immunoglobulin production and checked for the presence of CD58+/$CD23^{hi}$ Lymphoblast with three different panels (table 3.2) The aim of purity panel 1 is to check for presence of contaminants like monocyte, NK or, more important, T-Cells. Panel 1 markers are: CD3 BV785, CD20 AF488, CD38 BV711, IgD BV421, CD56 APC Cy7 (Biolegend); CD19 PerCP.Cy5.5, CD 27 APC (ebioscience); CD14 PE (BD).

The aim of panel 2 is to characterize the CD19+ B cells for their stage (mostly for the presence of lymphoblast CD38+ cells) and their production of immunoglobulins. Panel 2 markers are: CD3 PE Cy5 (Beckman coulter), CD19 PerCP.Cy5.5, CD 27 APC (ebioscience); CD20 AF488, CD38 BV711, IgM BV421, CD24 PE/Dazzle 594 (Biolegend); IgD PE, IgG APC H7 (BD Pharmigen™).

The aim of panel 3 is to check the cells for the presence of CD58+/CD23+ Lymphoblast. Panel 3 markers are CD19 PerCP.Cy5.5, CD 27 APC (ebioscience); CD20 AF488, CD38 BV711, CD23 BV421, CD58/LFA3 PECy5, CD14 BV510, CD56 BV510 (Biolegend), IgD PE (BD Pharmigen™), CD3 V500 (BD Horizon),

All cells have been checked for viability by means of Zombie-Acqua live staining (Biolegend) and fixed in IC fixation buffer (eBioscience).

Cell Sorting of B cells and FLS

At the end of the co-culture (initially between day 90 or 105; for later experiment 28 days has been chosen as the endpoint) cells were ready for FACS analysis. Last aliquot of total SN was collected from each well and stored as previously described. Versene (1:5000 EDTA (1x) Gibco) has

been chosen compared to trypsin in order to preserve surface markers. 150 µl p/w of versene have been added for 5 minutes at 37 degrees. This mild procedure usually results in the initial detachment of only B cells. After B-cell were collected in a 2 ml eppendorf tube, centrifuged for 10' at 1800 rpm, resuspended in cold PBS and kept in ice before staining procedure. Meanwhile other 150 µl of versene (this time cold) have been added for 20 min at room temperature in order to detach the FLS, which are more susceptible to apoptosis during detachment. After this step cells have been collected with the same procedure of the B cells.

a) B cell sorting:
B-cells have been stained using Panel 1 for characterization and sorted on CD19/CD20. Cells have been collected in a 1.5 ml eppendorf tube containing 1 ml of CM-B used immediately for DNA extraction. When number of cells was higher than expected, surplus cells have been frozen in freezing medium (90% CM-B, 10% DMSO).

b) FLS sorting:
FLS have been stained for GYPA, CD45, CD34,CD90 BV650 (BD Optibuild); CAD-11 PE, Podoplanin AF-647 (Biolegend). Cells have been sorted based on physical parameters and expression of CD90+. Then collected in a 1.5 ml eppendorf tube containing 1 ml of CM-FLS and finally stored at -80 or used immediately for RNA extraction.
Viability was assessed by means of Zombie NIR (Biolegend). Cells were not been fixed in order to perform molecular analysis.

FACS data analysis

All FACS analysis has been performed using the licensed software FLOWJO.

gDNA isolation and PCR for the BamH1 W repeat region of EBV (W-PCR) and calculation of frequency/level of EBV infected cells (Thorley-Lawson Laboratory):

For calculation of frequency of EBV infected cells a qDNA-qPCR with serial dilution has been developed in my lab following original Torgbor and Thorley Lawson Protocol [243]. As a positive control and a standard for the dilutions EBV+ Namalwa cells (ATCC CRL-1432) have been used. Compared to other cell lines, such as Nc-Nc - which at a single cell level contain a very variable number of genome copies (anywhere from 2-30) - Namalwa copies are integrated (2) and extremely more stable. Thanks to this feature, these cells proved to be a high-quality tool to interpolate the exact number of EBV-infected cells. EBV-Ramos (ECACC) cells have been used as a negative control.

a) Serial Dilutions and DNA extraction:

Dilutions were made as follows: $5x10^4$, 18'697 , 6'992, 2'614, 977, 365, 136, 51, 19, 7, 2, 1 (4 biological replicates for standard; 4 biological replicates for sample). Controls (B cell Time 0 and B- cell without fibroblasts) and sample have been placed in a 96 V bottom plate for subsequent PCR, spun at 1,200 rpm for 10 minutes at 4°C and the supernatant aspirated. To each well was then added 20 µl of digestion mix (prepared as follow:10X PCR buffer, 100 µl; Igepal (NP-40) 100 µl; Tween-20 100 µl; Proteinase K 50 µl of 20 mg/ ml stock and 650 µl of water). Plate was then sealed air-tight and incubated at 55°C overnight in an incubator. Then proteinase K deactivation was performed keeping the plate at 95°C for 10 minutes. Ten microliters (10 µl) of water were then added to all the sample and controls wells. b) W-repeat EBV gDNA-qPCR

For each reaction a master mix was prepared containing: 5 µl of Taqman Gene Expression Mastermix, 0.5 µl of 900 nM primers+250 nM fluorogenic probe, 2.5 µl of DEPcTreated Water and 2 µl of sample. Final reaction volume was 10 µl (See Table 3.4 for primer and probe sequences).

qPCR reactions were performed following Taqman recommendation for mastermix on an ABI 7900HT Fast Real-Time PCR System and a QuantStudio 5 Real-time PCR System: Step 1 (1 cycle): 2 min at 50°C; Step 2 (1 cycle) 10 min a 95°C; Step 3 (50 cycles): 15sec at 95°C, 1min at 60°C. Data were obtained with SDS and analysed using Prism. After W-repeat EBV gDNA qPCR plate has been collected and stored at -20 for electrophoresis and sequencing

Electrophoresis gel and Sanger Sequencing for EBV

a) Electrophoresis gel:

DNA concentration obtained from qPCR have been calculated using Quant-iT™ PicoGreen™ dsDNA Assay Kit and products have then used for electrophoresis analysis. Gel has been prepared using 4.5g of Agarose (Sigma Aldrich) diluted in 250 UltraPure TBE buffer 1X (ThermoFisher Scientific) to reach a percentage of 2, suitable to separate products between 200 base pairs (bp) and 2000 bp. 5 µl of products have been mixed with 1 µl of purple gel loading dye 6X (NewEngland Biolabs), loaded into the gel and run was performed in TBE buffer for 2 hours at 150 Volt. Gel results have been acquire using a Imaging System UV Transilluminator (UVP w/Sentech Stc-tb33usb). Gel products have been isolated and purified using a QIAquick Gel Extraction Kit (Quiagen) following Quiagen recommendation and prepared for sequencing.

b) Sanger Sequencing:

Gel products have been isolated and purified using a QIAquick Gel Extraction Kit (Quiagen) following Quiagen recommendation and concentration has been calculated using Nanodrop System (Thermofisher). To test integrity of products a 2nd round of PCR has been performed using "dummy" primers (unlabelled and not containing any probes). Best products have then been chosen and 5 µl have been sent for Sanger Sequencing using GATC biotech Service. Results have been visualized

using 4Peaks™ software and sequence alignment checked using Nucleotide Blast.

RNA purification, RT-PCR and cDNA qPCR for FLS:

RNA has been extracted from sorted FLS using Arcturus PicoPure™ RNA Isolation Kit (KIT0204). Elution has been performed with 12µl (3µl dead volume) resulting in a final volume of 9 µl. Such volume has been immediately retrotranscribed in a single reaction using a Superscript II Reverse trascriptase kit (Invitrogen). 23 µl of the sample volume obtained have been used in a pre-amplification step using Taqman PreAmp Master mix kit (final volume 50 µl) following Taqman recommendation for mastermix: Step 1 (1 cycle): 10 min a 95°C ; Step 2 (14 cycles): 15 sec at 95°C, 4 min at 60°C. At the same time other 5 µl of sample have been used to perform quantification using a Quant-iT™ PicoGreen™ dsDNA Assay Kit (Thermofisher). Final qPCR step has been performed in technical triplicates 10 microliters final volume: 5 µl of Taqman Gene Expression Mastermix, 0.5 µl of 900 nM primers+250 nM fluorogenic probe, 0.5 µl of DEPcTreated Water and 4 µl of sample. Primers used were: GAPDH (housekeeping), IL-6, IL-8 and TNF-a. Reactions were performed following Taqman recommendation for mastermix on an ABI 7900HT Fast Real-Time PCR System: Step 1 (1 cycle): 2 min at 50°C; Step 2 (1 cycle) 10 min a 95°C ; Step 3 (50 cycles): 15 sec at 95°C, 1min at 60°C. Data were obtained with SDS and analysed using Prism.

EBV PCR:

In order to identify EBV infection (CD19, EBNA-1, EBNA-2, LMP1, LMP2, EBER Table 3.5) PCR have been performed. RNA has been extracted and retrotranscribed has previously described. Reactions mix has been prepared as follow: 0.5µl of Q5® High-Fidelity DNA Polymerase, 2000U/ml (NEB), 30.5 µl of DEPc treated water, 10 µl of Q5 5X reaction

buffer, 0.5 µl of 10 µM Primer (F/R) 1 µl of 10 µM dNTPs and 5 µl of sample lysate or 500ng).: Step 1 (1 cycle): 2 min at 50°C; Step 2 (1 cycle) 10 min a 95°C ; Step 3 (40 cycles): 15 sec at 95°C, 1min at 60°C. Data were obtained with SDS. For this experiment Biorad CFX96 Real-Time PCR System have been used.

EBV Nested PCR:

In order to identify EBV infection (EBNA-2) 2 rounds of NESTED PCR have been performed. Nested PCR involves the use of two primer sets and two successive PCR reactions. RNA has been extracted and retrotranscribed has previously described. The first set of primers (Table 3.5) are designed to anneal to sequences upstream from the second set of primers and are used in an initial PCR reaction. For first round reaction mix has been prepared as follow: 0.5µl of Q5® High-Fidelity DNA Polymerase, 2000U/ml (NEB), 30.5 µl of DEPc treated water, 10 µl of Q5 5X reaction buffer, 1.5 µl of 10 µM Primer 1 EBNA 2F (sense), 1.5 µl of 10 µM Primer 1 EBNA 2I (antisense), 1 µl of 10 µM dNTPs and 5 µl of sample lysate or 500ng). For second round reaction mix has been prepared in the same way, with the exception of the primers: Primer 1 EBNA 2C (sense); Primer 2EBNA 2G/B (antisense). Cycling conditions for a total volume of 50 µL are the one suggested by NEB for Q5® (both 1st round and 2nd round). For this experiment Biorad CFX96 Real-Time PCR System have been used.

EBV Gradient PCR:

In order to identify EBV infection (EBER and LMP2A, Table 3.5) gradient PCR have been performed. Gradient PCR is a technique that allows the empirical determination of an optimal annealing temperature. RNA has been extracted and retrotranscribed has previously described. Reaction mix

has been prepared in a total volume of 10 µl as follow: 0.5 µl of 10 µM of both forward and reverse primers (LMP2A/EBER), final concentration 500nM/L, 3 µL of DEPc water, 5 µM of SYBR green blue mix (qPCRBIO) and 1 µl of templated (5ng/ µL). Cycling condition: Step 1 (1 cycle): 2 min at 50°C; Step 2 (1 cycle) 15 min a 95°C ; Step 3 (50 cycles): 15 sec at 95°C, 1min at 70°,69.5°,68.4°,66.4°,64.0°,62.0°,60.7°,60°C. Data were obtained with SDS and analysed using Prism 50. For this experiment Biorad CFX96 Real-Time PCR System have been used.

Enzyme-linked immunosorbent assays (ELISA) and Custom LEGENDplexTM Assay Panel test on supernatant from co-culture:

All supernatants collected every 14 days from each well of the co-culture system have then been tested on several different ELISA and multiplex assay from biolegend in order to asses any variation in term of immunomodulation. Test have been performed for: total IgG, IgM and IgA (Bethyl Laboratories); IgG, IgM and IgA ACPA (Axis-Shield); Viral protein IgG-EBNA, IgG-VCA (Virion/Serion); IL-6, IL8 (ELISA MAX, Biolegend); April (ebioscience); BAFF/TNSF13B (R&D). LEGENDplexTM have been performed following company instructions. Analytes tested are APRIL, BAFF, CCL2 (MCP-1), CCL3 (MIP-1α), CXCL10 (IP-10), IL-6, IL-12p70, IL-22, IL-27, RANKL, TNF-α, TSLP, Rantes.

Single cell sorting, digestion, RT-PCR and BamHIqPCR for EBNA-LP transcript:

Single cell sorting, digestion and RT-PCR has been performed as previously described by Corsiero [244]. qPCR reactions were performed following Taqman recommendation for mastermix on a QuantStudio 5 Real-time PCR System: Step 1 (1 cycle): 2 min at 50°C; Step 2 (1 cycle)

10 min a 95°C; Step 3 (50 cycles): 15sec at 95°C, 1min at 60°C. Data were obtained with SDS and analysed using Prism. Data have bene normalized on housekeeping gene GAPDH.

Co-culture and Singles cells 10X genomics chromium sequencing:

Sorted EBV+ RA-CD19+ cells at end of a normal co-culture assay have been plated again on RA-FLS, without stimuli. For each condition 1 wells from a 12well plate with 30'000 FLS and 93'000 B cells (Number has given by the Time 0 Control "leftover"; these cells were later dead and have not been use for this experiment). Cells have been cultured in 2ml of CM-B cells for 14 days. Medium has been changed and collected every 7 days.

At the endpoint, all medium has been collected gently to not disturb the RACD19+cell/RA-FLS interaction. Cells have been detached with 5 min trypsin at 37°C, collected in a 5ml eppendorf tube, spin down at 1500rpm for 5 minutes, resuspended in 100µl and kept on ice during transfer to the Genome Center of Queen Mary University (QMUL). Upon arrival, viability and number of cells was assessed using a Countess II automated cell counter; concentration trough centrifuge was operated when necessary. Preparation of sample and next steps have been performed following 10X suggestion:1000 cells have been loaded on the 10XChromium Controller instrument with 10x genomic reagents for VDJ and 5'transcriptomic analysis in order generate the libraries. Sequencing have been obtained with Illumina sequencing system.

(Loading on Chromium Controller instrument and Illumina Sequencing have been performed by Genome Centre QMUL personnel). Output data obtained by Illumina have been analysed using free 10X software LOUPE and VLOUPE.

Alternative analysis of Carejavi and comparison with other EBV infected cell line have been performed using R studio for Mac (Version 1.3.1093) and Seurat 3.0 from SatijLab [245] [246]. Codes used for quality control

(QC) and analysis have been obtained from SatijLab webpage (https://satijalab.org/seurat/). Cell line used in comparison (GM12878 and GM18502) have been obtained from Osorio et al [247], are deposited in the NCBI Sequence Read Archive (SRA) and can be accessed along with processed data from the NCBI Gene Expression Omnibus (GEO, Series Accession: GSM484896 (for GM12878) and GSM2392689 (for GM18502)). In QC, cells having i) a number of features between 200 and 6000, ii) a value less than 10% in mitochondrial genes, have been subset for analysis.

Search Tool for the Retrieval of Interacting Genes/protein (STRING) Analysis:

STRING is a biological database of known predicate protein/protein interaction. This database contains numerous information about these kinds of interaction, like experimental data, computational prediction methods and public text collection. For more information, visit STRING website online [248].

Statistical analysis:

Differences in quantitative variables were analysed by Mann Whitney U-test or Kruskal-Wallis with Dunn's post-test when comparing two or multiple groups, respectively. To test if the values were coming from a Gaussian distribution the Shapiro-Wilk normality test has been ran.
For 10X, Cell ranger build-in software has run Principal component analysis (PCA), t-distributed Stochastic Neighbor Embedding (t-SNE) , clustering and differential expression: from 10X genomic: *For each feature three values per cluster are computed: 1)The mean Unique molecular identifiers (UMI) counts per cell of this feature in cluster 2)The log2 fold-change of this feature's expression in cluster i relative to all other clusters 3)The p-value denoting significance of this feature's expression in cluster i*

relative to other clusters, adjusted to account for the number of hypotheses (i.e. number of features) being tested.

All statistical analyses were performed using GraphPad Prism version 6, (GraphPad Software, San Diego Ca, USA); Cell Ranger, LOUPE, VLOUPE (10Xgenomics); STRING. A p value <0.05 was considered statistically significant.

CHAPTER 4 TECHNICAL RESULTS

In order to obtain long-standing EBV-infected B cells, I had to setup an *in vitro* co-culture system (Fig.4.1) whereby RA-FLS obtained from synovial fluid and synovial tissue of ACPA+ patients could have been cultured with B cells obtained from ACPA+ RA or healthy donor, with the final aim of understanding the long-term interactions between these two cells.

Preliminary experiments and arrangement of co-culture system.

Creating a biobank

First, I needed to collect enough samples to obtain preliminary data. Thanks to the availability in my department, during the first months of my PhD has been possible to collect samples of different kind (synovial fluid, synovial tissue and peripheral blood) from patients (RA, osteoarthritis (OA), psoriatic arthritis (PsA)) and healthy donor. Thus, during these four years I have been able to create a small biobank of FLS, PBMCs and B cells that has allowed me not just to carry over my PhD but also to collaborate in my lab to several projects not directly related to EBV [249].
Also, thanks to my supervisor Prof. Michele Bombardieri, I managed to obtain an official phlebotomist training from the National Health Service (NHS), in order to draw blood. This has not just allowed me to work independently when it comes to blood collection but also has gave me an incredible opportunity to directly interact with the patients.

A list of samples used in this project is attached in table 4.1 and 4.2 in Appendix

Successful culture of FLS from synovial tissue and synovial fluid

During isolation, FLSs started to emerge from both synovial fluid and synovial tissue samples. The emerging cells showed an undoubtable fibroblast shape, were clearly alive when stained for trypan blue and have been dividing and proliferating at different rates. Most importantly, FLS from synovial fluid did not show any strong macroscopic differences when compared to their tissue counterpart.
FLS were isolated and cultured successfully and have been frozen after passage 3 and beyond. All the experiments have been performed on frozen FLS for convenience: long time to reach passages 4 or 5, and timing of blood samples. Figure 4.1 is showing the 040/2016 RA-FLS, one of the most used sample, prior to the first co-culture experiment with total PBMCs: cells were alive and at a confluence of 80%.

First co-cultures arrangement: several negative results

The set-up of the co-culture system has required numerous attempts before demonstrating to be reliable. In this section I will quickly go through some of these attempts, but I will not discuss the data. In my first co-culture arrangement I decided to use total PBMCs obtained both from RA and HD patient. I choose a 96w/ell plate using 1,5 x 10^5 PBMCs por well. Once combined with the RA-FLS, the cells where cultured in (CM-B Cells) in 250 µl (CM-B Cells) with stimuli. As control I used FLS and PBMCs alone, with or without stimuli. At the time I had two endpoints for the experiment, 14 and 28 days, and mostly I wanted to see any sort of interaction between the cells, the production of EBV related antibodies (EBNA IgG) and evidences of the presence of the virus. The final operation would have been extracting the RNA directly from the mix of cells, retrotranscribe it to cDNA, in order to perform sybr green qPCR using a group of primers available in my lab and designed by Dr. Croia.

(table 3.5). These primers were supposed to target specific B cells genes (such as CD19) or viral sequences (such as EBER or EBNA-1).

After plating the cells, the co-culture was not checked day by day under the microscope to track the evolution of the experiment and I just decided to extract the total RNA after 14 and 28 days. At these two endpoints I also collected the SN in order to test it for the EBNA IgG. Unfortunately, after 4 attempts the RNA yield tested at the Nanodrop was almost null. A preamplification step was used after the RT-PCR in order to increase the chances for a signal, but no amplification was obtained (data not shown). Also, I did not identify any product in the agarose gel 2% (data not shown).

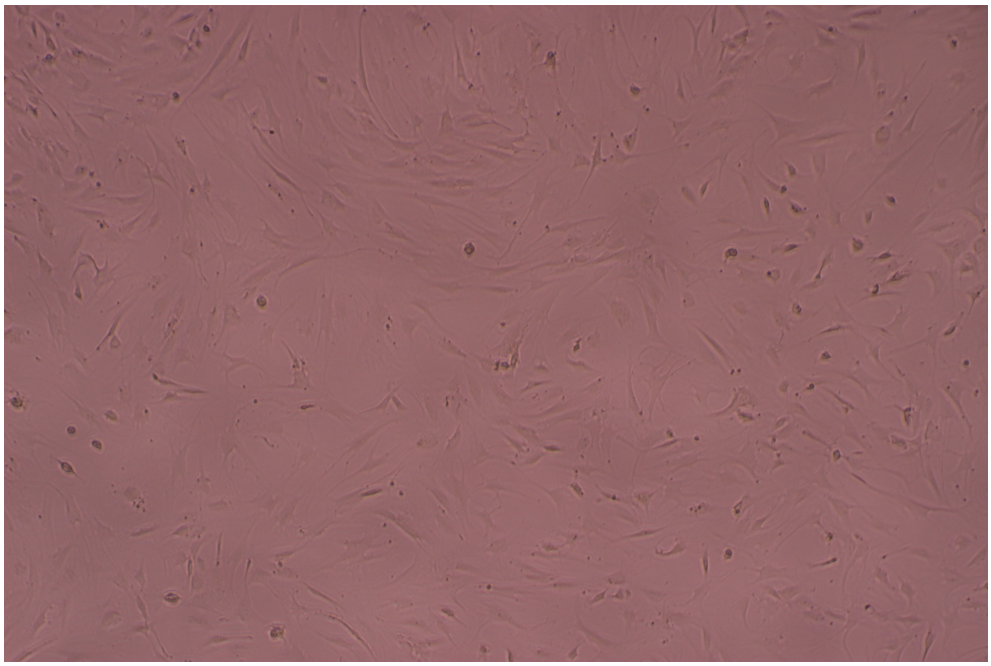


Figure 4.1. RA-FLS 040/2016 prior to co - culture (2016). Photo acquired with Olympus Microscope (brightfield).

Furthermore, all the SN tested on ELISA for EBNA IgG resulted negative or very close to the negative control (data not shown)

From these first results it looked almost like no cells were present in the co-culture, neither after 14 days nor 28 days. I decided then to repeat the experiment and to check the co-culture of PBMCs and FLS under the microscope in order to track the evolution and to identify any issue.

After performing this simple operation, the problem was quite clear: the PBMCs where showing a strong immune response toward the FLS in culture due to MHC mismatch (probably due the fraction in monocytes/macrophages or T CD8+ killer cells). At different degrees of level between the patients, the immune cells were responding to the FLS as foreign elements and this was due to heterologous nature of the assay. Indeed, obtaining FLS and PBMCs from the same patient has so far resulted quite difficult and was possible only in one case, as previously stated. From several picture obtained was evident that the PBMCs started to target the FLS already in the first hours during the experiments. After 14 days most of the FLS were gone or showed ongoing apoptotic process

while after 28 days most of PBMCs left were strongly positive for trypan blue (data not shown).

Get a clearer shot: 2nd co-culture optimization

CD19+ positive beads purification instead of total PBMCs

For the purpose of reducing the variability – thus the kind of cells - inside the experiment, I decided to design a more controlled system. The first main change applied to the co-culture has been to perform a CD19+ positive selection, in order to get rid of any possible "contamination" from T cells or macrophages. Apart from the main concern above mentioned (the immune response toward FLS), the other issue I wanted to solve was to eliminate the T cells or the APC from the co-culture. Indeed, I wanted to confirm previous data produced from Kam and Bombardieri in my lab, showing that the FLS are able to some extent to trigger SHM and CSR[126].
This has resulted in a clearer arrangement for my purposes, as can be seen by microscope by figure 4.3 and figure 4.4.

Increased amount of medium

Due to the long-lasting nature of the experiment, the viability of the cells can be easily affected. One solution would have been to change the media, but unfortunately, with the 96 well plate assay this was resulting in either detaching the B cells from the FLS or losing all the floating cells. In order to solve this issue, I decide to switch for a 48 well plate, increase the amount of medium to 1.2 ml in each well (from previous 250 µl) and to collect and change the medium only every 14-15 days. This have made easier the medium collection step at any time point, without stressing the cells (Fig 4.2).

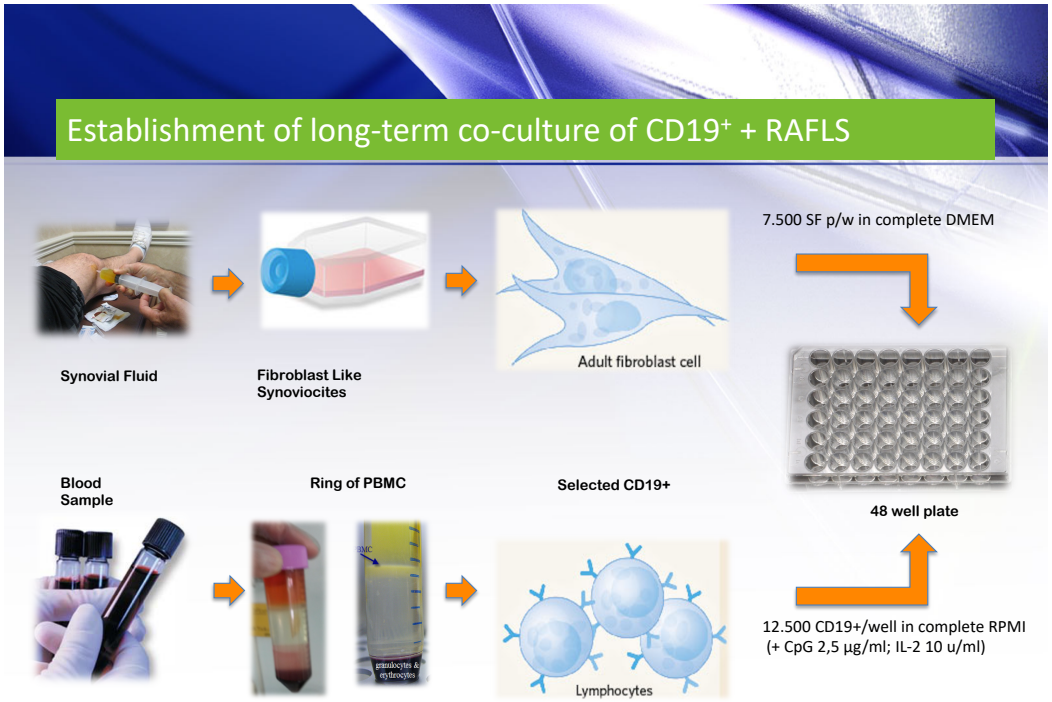


Figure 4.2. Summary of the co-culture 2n[d] setup. Synovial fluid RA-FLS are co-cultured with CD19+ cells in a 48 well plate. CpG and Il-2 are added as stimuli.

Co-culture system allows both FLS and CD19+ to survive up to 90 days

The 2nd set-up of the co-culture system has shown through microscopy what would be later confirmed by FACs analysis (see next chapter): a high percentage of both FLS and CD19+ cells were still alive at the end of the experiments compared to control (Fig 4.4)
This has been indirectly confirmed also by ELISA in terms of immunoglobulins and cytokines production. Furthermore 3 out of 9 of the preliminary co-cultures during the set-up phase showed live cells still at 105 days. This means that these cells have been able to survive in vitro for more than 3 months and this is a remarkable survival rate for B cells which are known to be usually short lived in the absence of strong survival signals. It is important to mention that the co-culture experiments were not run in parallel in order to allow a thorough examination at the endpoint. At the time a reliable system to detect and quantify EBV was not in place, so these cultures have not been screened for the virus.

RA-FLS induce class switched antibodies production from RACD19+ cells

Preliminary data have shown that production of total IgM, IgG, and IgA from the CD19+ B cells co-cultured with RA-FLS under IL-2 and CpG stimulation is greatly increased compared to controls (Fig 4.5 & 4.6). Also, the ratio between IgM and IgG seems to vary along the time, as another confirmation of class switching. This has been shown in all my experiments with a slight variation between isotypes among different co-culture. It is important to highlight that this is a highly specific characteristic of RA-FLS since dermal fibroblasts obtained from the same donor failed to induce Ig release[126]. However, I was not able to detect

antibodies against EBNA-1-IgG, VCA and CCP antigens (including ACPA-IgG, ACPA-IgM, ACPA-IgA HCP1 and 2, VCP1 and 2 in any of the co-culture supernatants (SN) (data not shown).

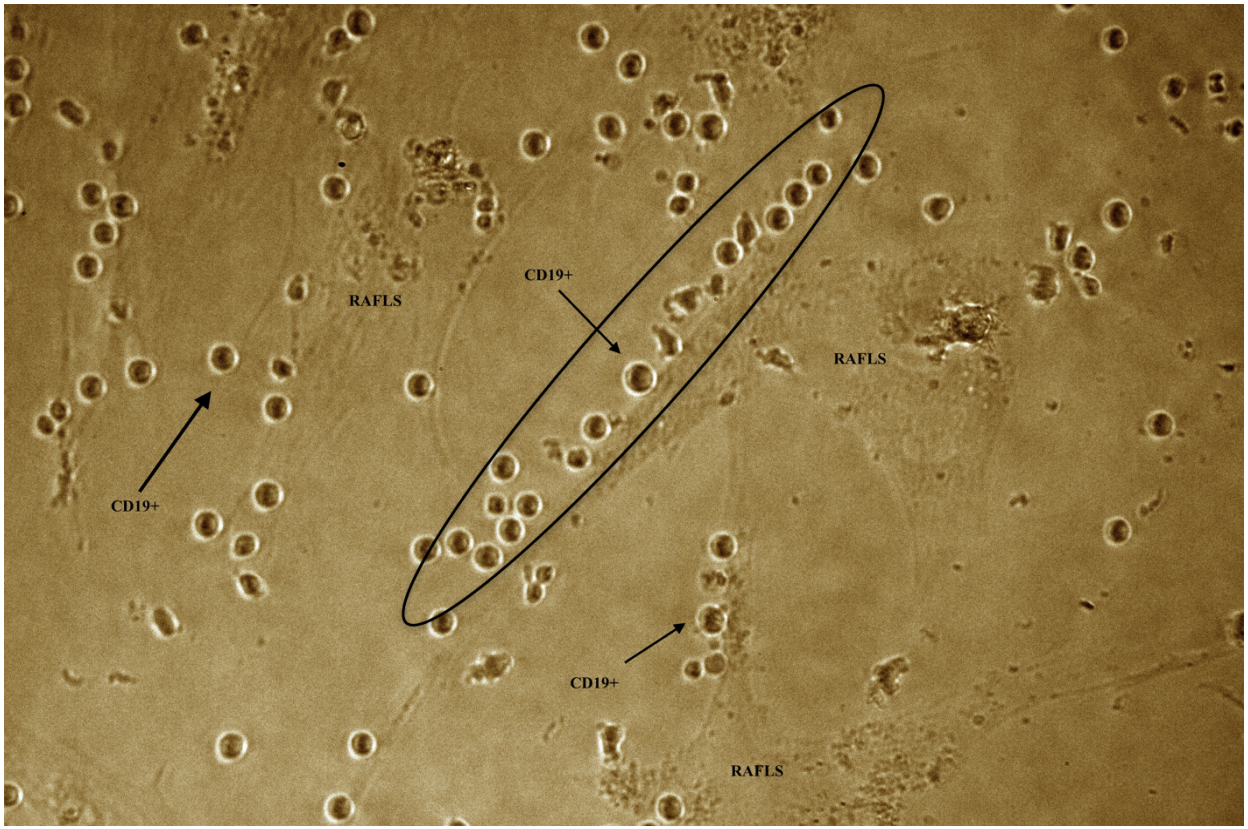


Figure 4.3. RA-FLS and isolated cells in co-culture on day 28. Isolated cells are presumed to be CD19+ obtained after purification. Photo acquired with Olympus Microscope (bright field).

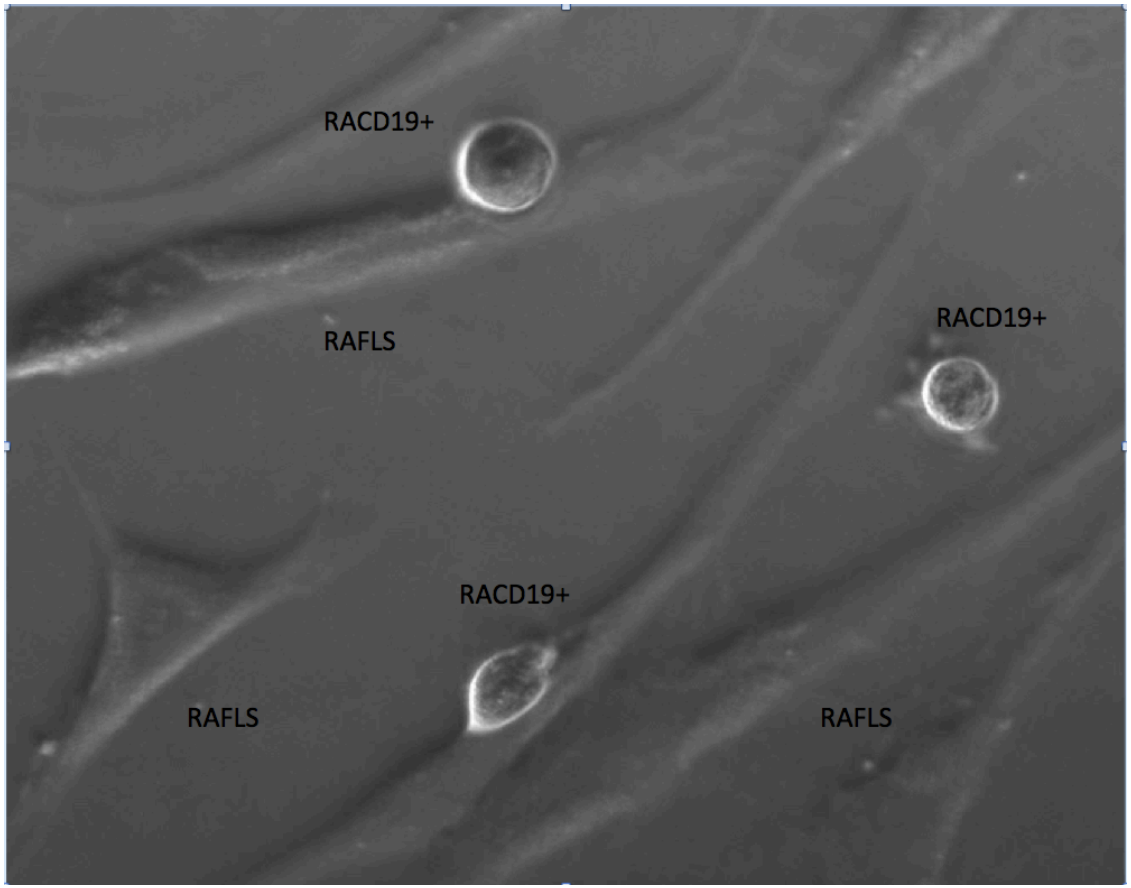
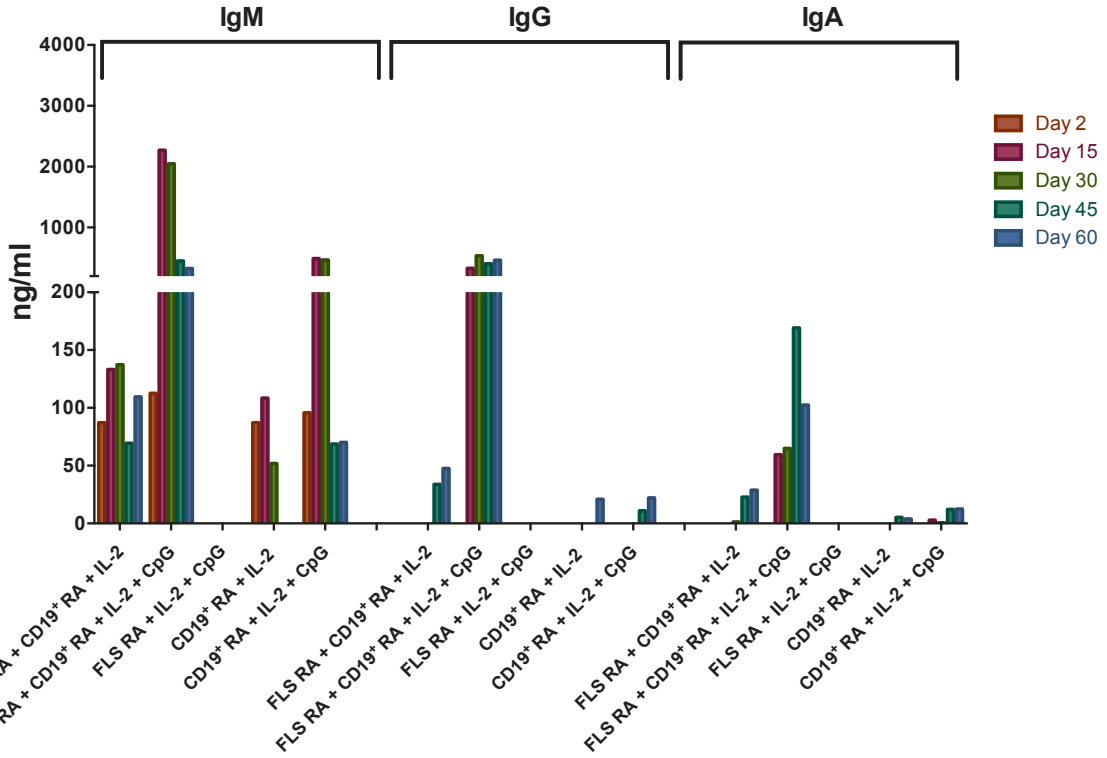


Figure 4.4. RA-FLS and isolated cells in co-culture on day 90. Isolated cells are presumed to be CD19+ obtained after purification. Photo acquired with Olympus Microscope (brightfield).


IgM
IgG
IgA
4000
3000
2000
1000
200
150
100
50
0
ng/ml
Day 2
Day 15
Day 30
Day 45
Day 60
FLS RA + CD19+ RA + IL-2
FLS RA + CD19+ RA + IL-2 + CpG
FLS RA + IL-2 + CpG
CD19+ RA + IL-2
CD19+ RA + IL-2 + CpG

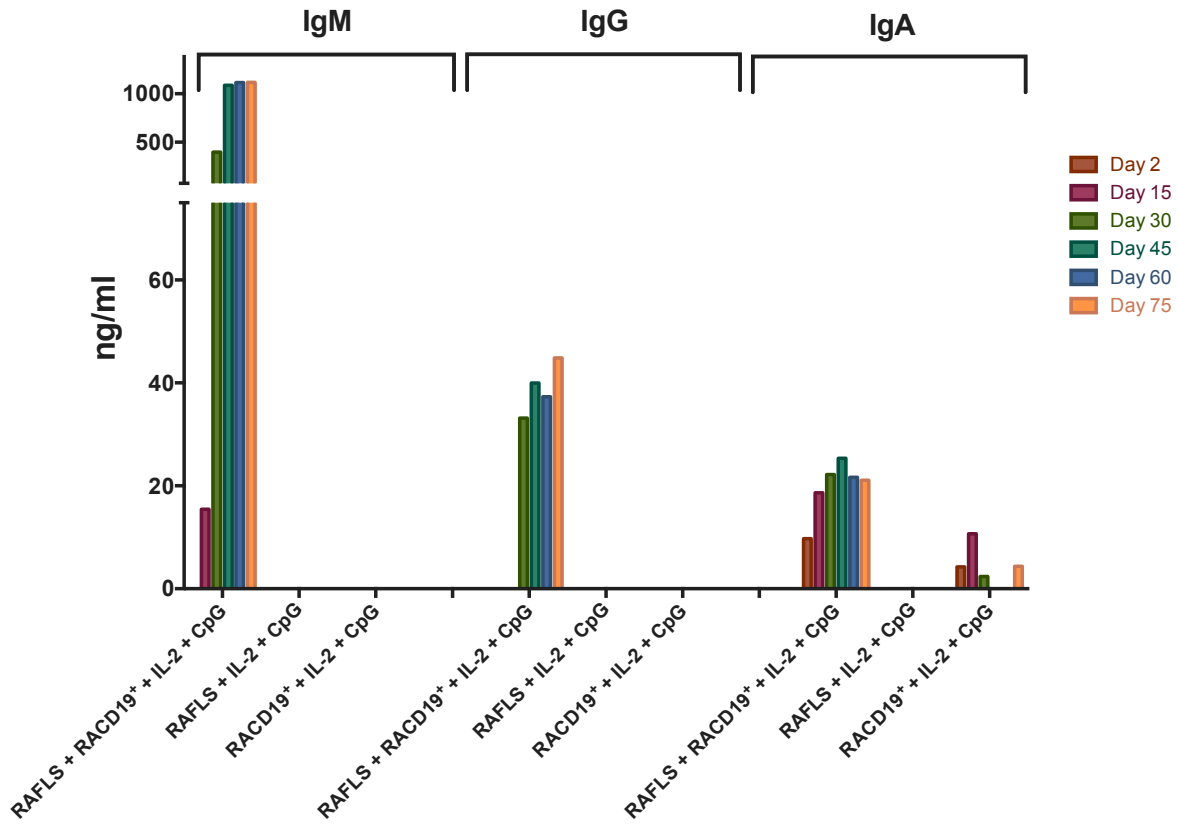
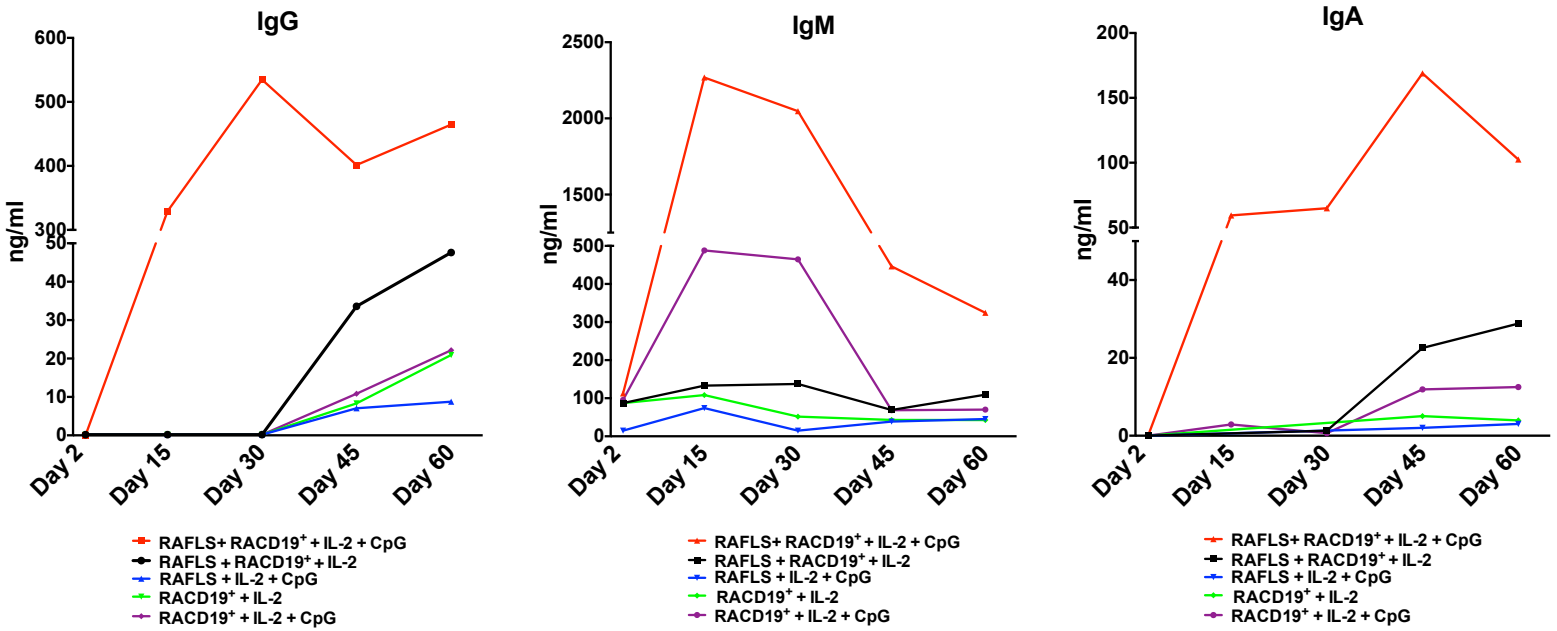


Figure 4.5. ELISA results for total IgM, IgG and IgA concentration (starting from previous page). IgM, IgG and IgA (B) concentration from 2 different co-culture of RA-FLS and RACD19+ (both showing sign of proliferation) cells from day 2 to day 60/75.


IgG
ng/ml
600
500
400
300
50
40
30
20
10
0
Day 2
Day 15
Day 30
Day 45
Day 60
RAFLS+ RACD19+ + IL-2 + CpG
RAFLS + RACD19+ + IL-2
RAFLS + IL-2 + CpG
RACD19+ + IL-2
RACD19+ + IL-2 + CpG
IgM
ng/ml
2500
2000
1500
500
400
300
200
100
0
Day 2
Day 15
Day 30
Day 45
Day 60
RAFLS+ RACD19+ + IL-2 + CpG
RAFLS + RACD19+ + IL-2
RAFLS + IL-2 + CpG
RACD19+ + IL-2
RACD19+ + IL-2 + CpG
IgA
ng/ml
200
150
100
50
40
20
0
Day 2
Day 15
Day 30
Day 45
Day 60
RAFLS+ RACD19+ + IL-2 + CpG
RAFLS + RACD19+ + IL-2
RAFLS + IL-2 + CpG
RACD19+ + IL-2
RACD19+ + IL-2 + CpG

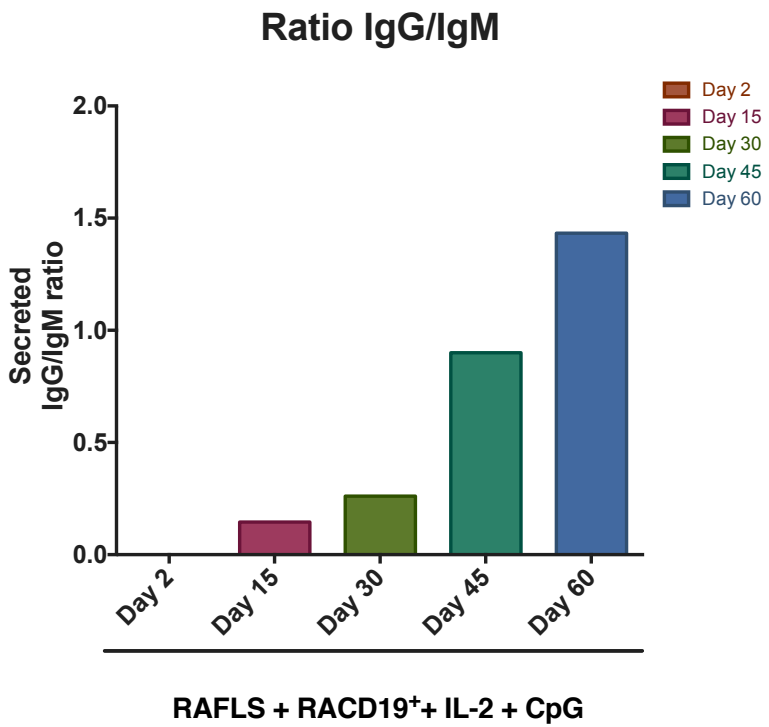


Figure 4.6. IgG, IgM, IGA production at day 2-60. Ration between IgG and IgM between day 2 and day 60 (starting from previous page). IgG, IgM and IgA production from day 2 to day 60 from CD19+ co-cultured with RA-FLS plus stimuli (CpG + Il-2) is higher than control. Same results are merged to be expressed in term of IgG/IgM ratio.

It is possible that if ACPA+ EBV-infected B cells were present and proliferating in the co-culture, the level of anti-citrullinated antibodies might be below the level of detectability of the ELISA test used. Attempts to concentrate IgG in supernatant have been performed but failed. Alternative approaches (WB, ELISA for specific citrullinated antigens etc) are going to be taken into account.

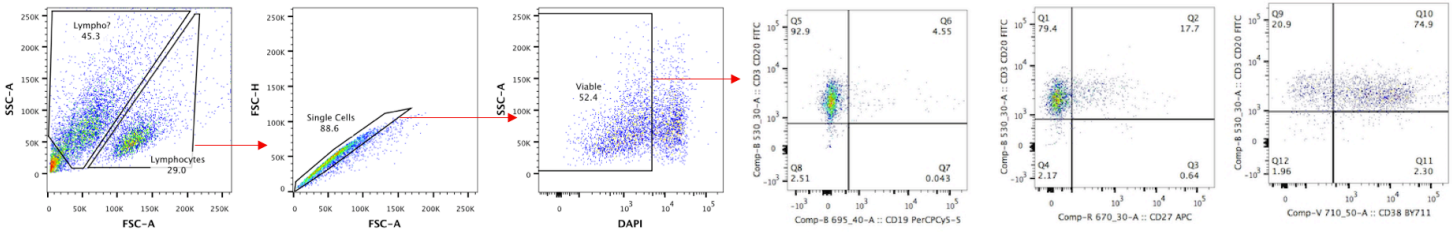
The introduction of flowcytometry and cell-sorting to assess purity and characterize the sample before and after co-culture

First panel design and sorting for co-culture RA-FLS and RACD19+: T cell contamination issue

In order to analyse the evolution of the different subsets of B cells inside of the co-culture and to obtain a higher level of purity during molecular biology screening, I decided to adopt flowcytometry and single cells sorting analysis.

The first small panel I designed contained only a few marker (DAPI, CD19, CD20, CD3, CD27 and CD38, see table 4.3 and methodologies) and was tested on the co-culture of sample RA-FLS 034/2016 and RACD19+099/2016, both ACPA+, at the end of the test. In order to increase cell purity, during the acquisition of the panel I decided to set up two fluorescence minus one staining, or FMO (CD20 and CD3). Unfortunately, this panel was designed with several bias – probably due to

lack of experience – since instead of aiming for a double positive population of CD19 and CD20 I only targeted the CD19 cells. Also, CD20 and CD3 – a common T cell human marker - were occupying the same fluorochrome (FITC). Nevertheless, this panel have revealed that - despite the CD19+ purification - the amount of T cells in the co-culture was still relevant (Figure 4.7).


Lympho?
45.3
Lymphocytes
29.0
SSC-A
FSC-A
Single Cells
88.6
FSC-H
FSC-A
Viable
52.4
SSC-A
DAPI
Q5
92.9
Q6
4.55
Q8
2.51
Q7
0.043
Comp-B 530_30-A :: CD3 CD20 FITC
Comp-B 695_40-A :: CD19 PerCPCy5-5
Q1
79.4
Q2
17.7
Q4
2.17
Q3
0.64
Comp-R 670_30-A :: CD27 APC
Q9
20.9
Q10
74.9
Q12
1.96
Q11
2.30
Comp-V 710_50-A :: CD38 BY711

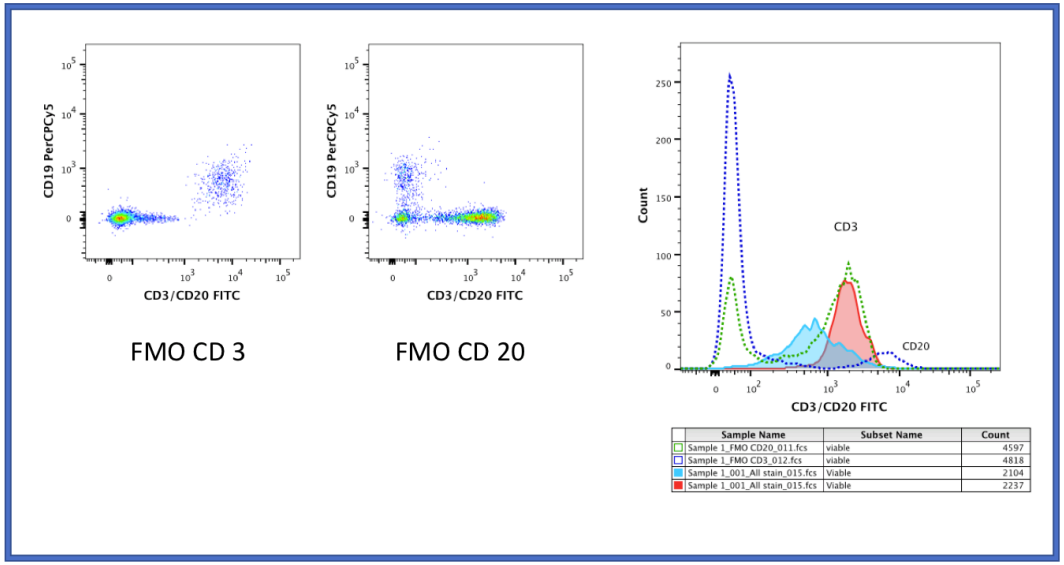
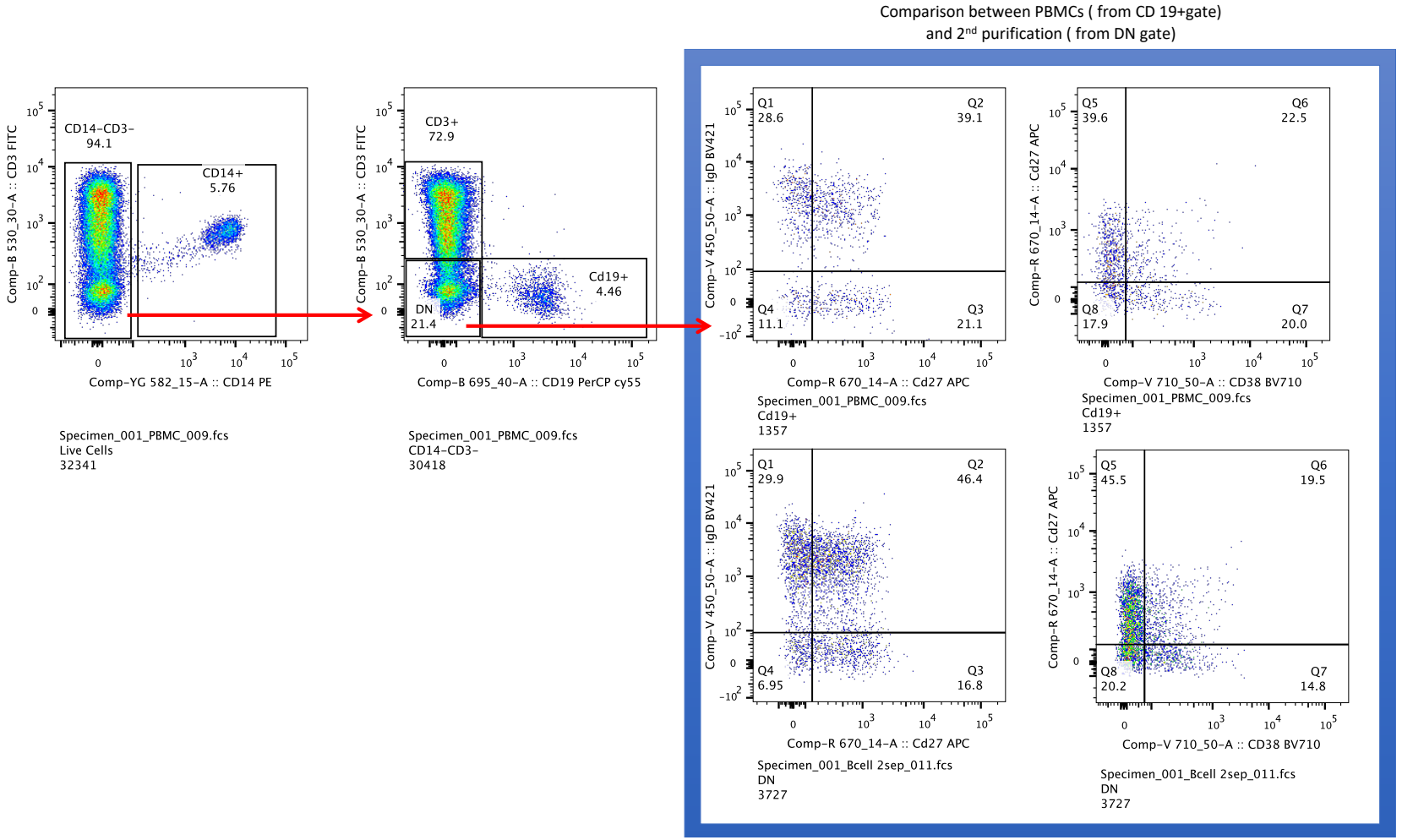

CD19 PerCPCy5
CD3/CD20 FITC
FMO CD 3
CD19 PerCPCy5
CD3/CD20 FITC
FMO CD 20
Count
CD3
CD20
CD3/CD20 FITC
Sample Name | Subset Name | Count
Sample 1_FMO CD20_011.fcs | viable | 4597
Sample 1_FMO CD3_012.fcs | viable | 4818
Sample 1_001_All stain_015.fcs | Viable | 2104
Sample 1_001_All stain_015.fcs | Viable | 2237

Figure 4.7. First FACS analysis and T cells contamination issue (previous page): First panel design and acquisition of co-culture RA-FLS 034/2016 and RACD19+099/2016. T cell contamination in the co-culture after 28 days can be identified through the CD20 FMO. Cells have been gated on physical parameters, then singlets, then DAPI- and finally on CD3_CD20/CD19, CD3_CD20/CD27, CD3_CD20/CD38.

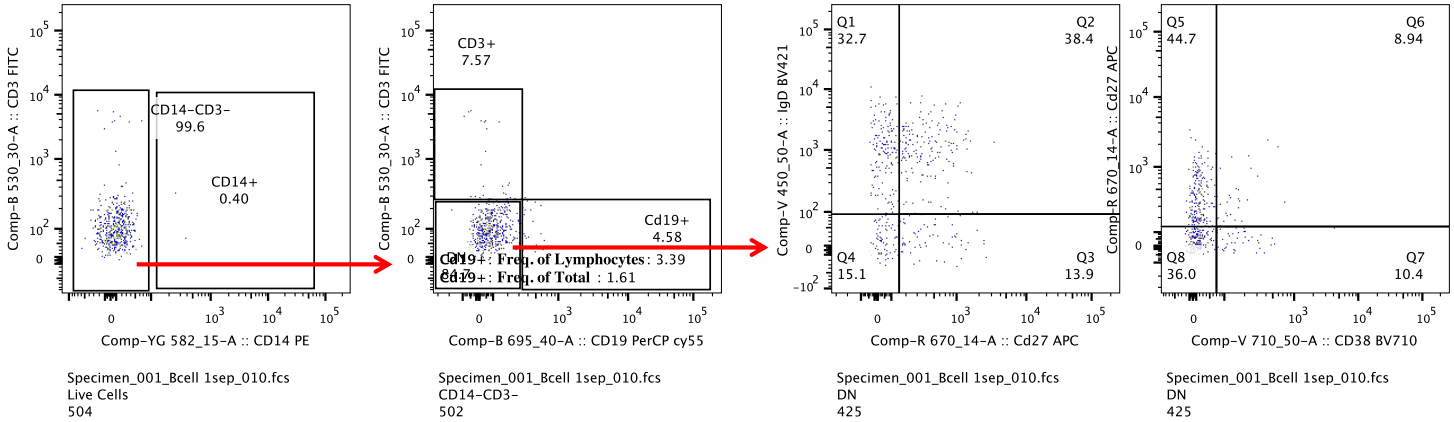
Purification yield increased after 2nd passage on column and masked antigen issue solved

In order to increase purity, I decided - during the purification steps - to pass the CD19+ enriched fraction a 2nd time on a 2nd column and test cells with a second panel: Zombie Aqua (instead of DAPI) CD3, CD14 (for monocytes) and then typical B cells markers such as CD19, IgD and CD27 (but not CD20) (see table 4.3). In this case, I used blood collected from a healthy donor. While the T cells issue appeared to be solved (T cells fraction dropped from 72.9% in total PBMCS to 5.67 % in the 2nd B cell separation), the majority of cell in the CD3/CD19 appeared as double negative (DN, Figure 4.8). Nevertheless, this DN population showed the same pattern for IgD and CD27 when compared to the CD19+ cells. This made me speculate that the CD19+ purification antibody might have targeted the same antigen as the FACs one, thus masking the epitope. A 3rd panel to address this has been created, adding this time CD20 and testing it again on a healthy donor sample before and after separation. I also added CD56 as a further marker for NK cells identification. This experiment resulted in obtaining CD20+ cells (81% after 2nd separation) and a small CD19+CD20+ fraction (18.8%) compared to the CD19+CD20+ population present in the PBMCs (28.3%) (Fig 4.9).
Also, T cells exclusion worked more efficiently resulting in a drop from 68.0% (PBMCs) to a 4.62% (1st separation) and to 0.27% (2nd separation).
It is important to mention that B cells negative selection has been taken into account, but T contamination percentage was higher than in the 1st positive separation (data not shown).

PBMCs


Comparison between PBMCs (from CD 19+gate)
and 2nd purification (from DN gate)
CD14-CD3-
94.1
CD14+
5.76
Comp-B 530_30-A :: CD3 FITC
Comp-YG 582_15-A :: CD14 PE
Specimen_001_PBMC_009.fcs
Live Cells
32341
CD3+
72.9
DN
21.4
Cd19+
4.46
Comp-B 695_40-A :: CD19 PerCP cy55
Specimen_001_PBMC_009.fcs
CD14-CD3-
30418
Q1
28.6
Q2
39.1
Q4
11.1
Q3
21.1
Comp-V 450_50-A :: IgD BV421
Comp-R 670_14-A :: Cd27 APC
Specimen_001_PBMC_009.fcs
Cd19+
1357
Q5
39.6
Q6
22.5
Q8
17.9
Q7
20.0
Comp-R 670_14-A :: Cd27 APC
Comp-V 710_50-A :: CD38 BV710
Specimen_001_PBMC_009.fcs
Cd19+
1357
Q1
29.9
Q2
46.4
Q4
6.95
Q3
16.8
Comp-V 450_50-A :: IgD BV421
Comp-R 670_14-A :: Cd27 APC
Specimen_001_Bcell 2sep_011.fcs
DN
3727
Q5
45.5
Q6
19.5
Q8
20.2
Q7
14.8
Comp-R 670_14-A :: Cd27 APC
Comp-V 710_50-A :: CD38 BV710
Specimen_001_Bcell 2sep_011.fcs
DN
3727

CD19+ 1st Passage



CD19+ 2nd Passage

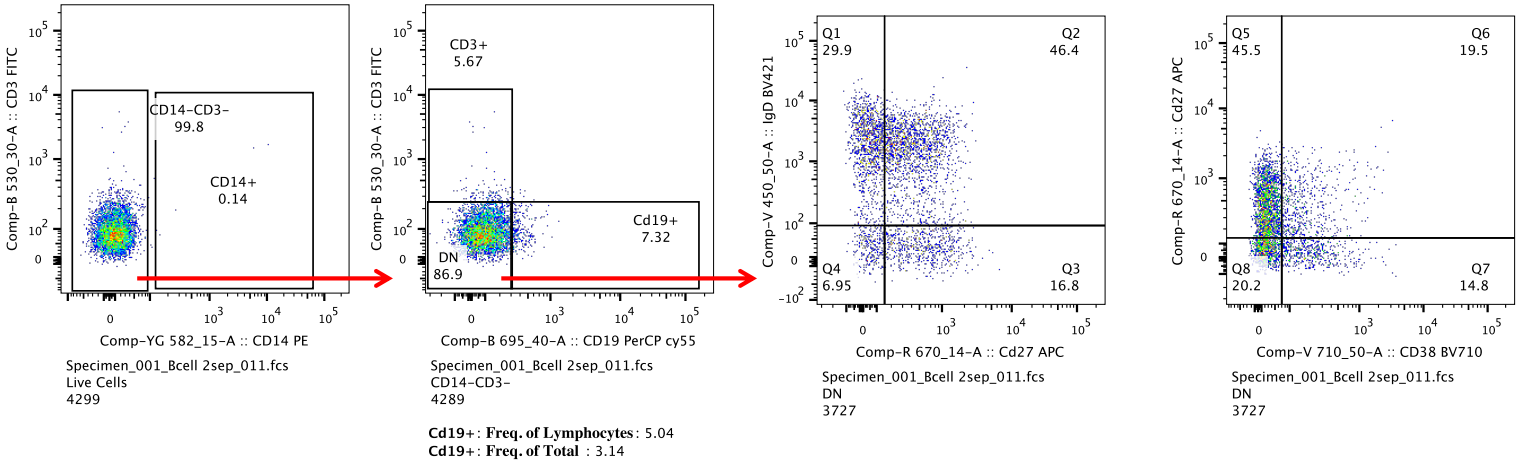
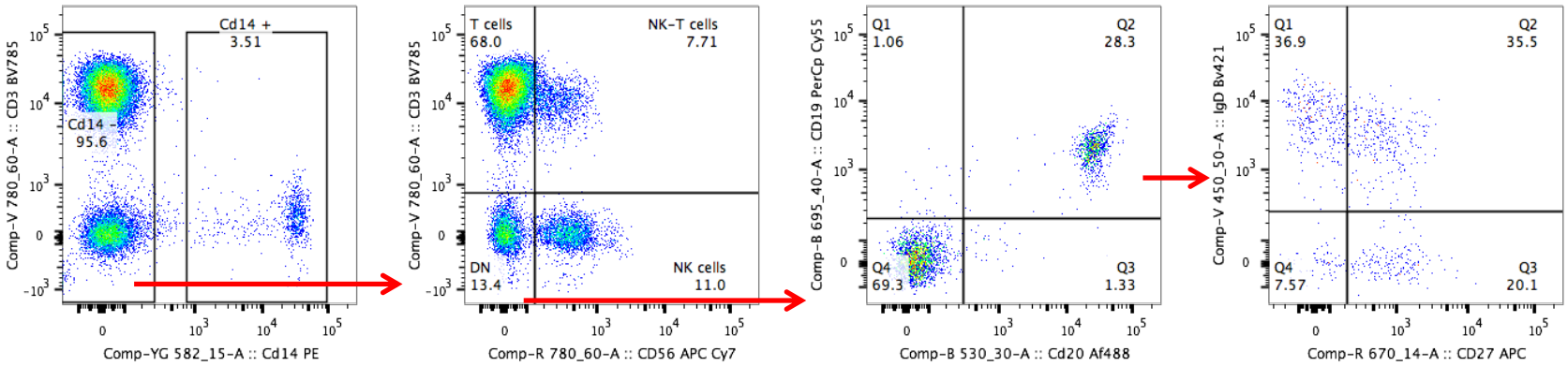
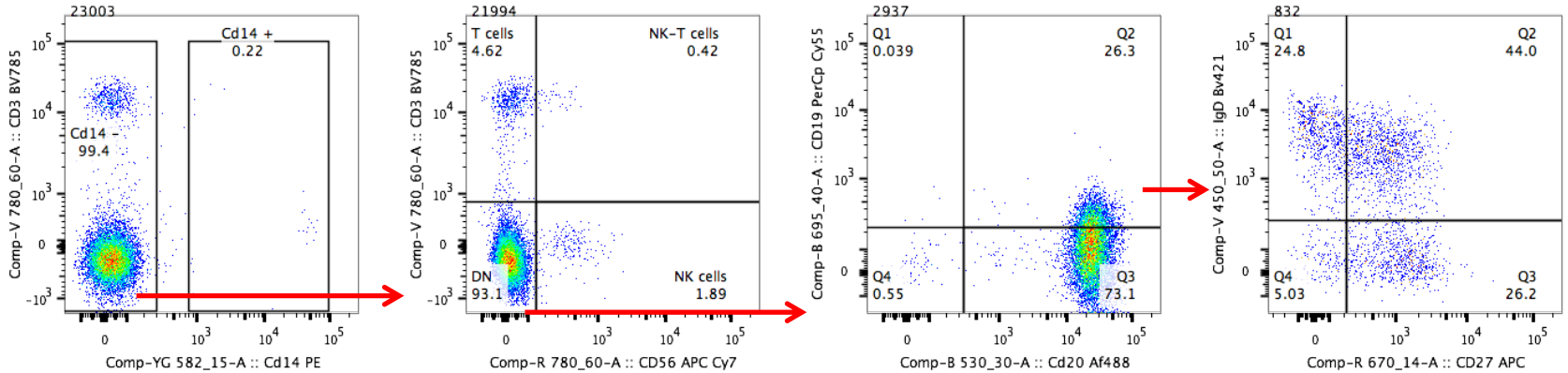


Figure 4.8. 2nd FACS Panel, T cells contamination solved and double negative issue (previous pages, first figure is showing total PBMCs while second figure is showing 1st separation and 2nd separation) . With the 2nd FACs panel I managed to solve the T cells contamination problem. Nevertheless, in CD3/CD20 graph most of the cells resulted double negative. This double negative population appeared almost identical in term of IgD and CD27 expression when compared to the CD19+ cells. Cells have been gated on physical parameters, then singlet, Aqua-, CD14-, CD19+ or DN.

PBMCs


Cd14 +
3.51
Cd14 -
95.6
Comp-V 780_60-A :: CD3 BV785
Comp-YG 582_15-A :: Cd14 PE
T cells
68.0
NK-T cells
7.71
DN
13.4
NK cells
11.0
Comp-R 780_60-A :: CD56 APC Cy7
Q1
1.06
Q2
28.3
Q4
69.3
Q3
1.33
Comp-B 695_40-A :: CD19 PerCp Cy55
Comp-B 530_30-A :: Cd20 Af488
Q1
36.9
Q2
35.5
Q4
7.57
Q3
20.1
Comp-V 450_50-A :: IgD Bv421
Comp-R 670_14-A :: CD27 APC

1st Passage


23003
Cd14 +
0.22
Cd14 -
99.4
Comp-V 780_60-A :: CD3 BV785
Comp-YG 582_15-A :: Cd14 PE
21994
T cells
4.62
NK-T cells
0.42
DN
93.1
NK cells
1.89
Comp-R 780_60-A :: CD56 APC Cy7
2937
Q1
0.039
Q2
26.3
Q4
0.55
Q3
73.1
Comp-B 695_40-A :: CD19 PerCp Cy55
Comp-B 530_30-A :: Cd20 Af488
832
Q1
24.8
Q2
44.0
Q4
5.03
Q3
26.2
Comp-V 450_50-A :: IgD Bv421
Comp-R 670_14-A :: CD27 APC

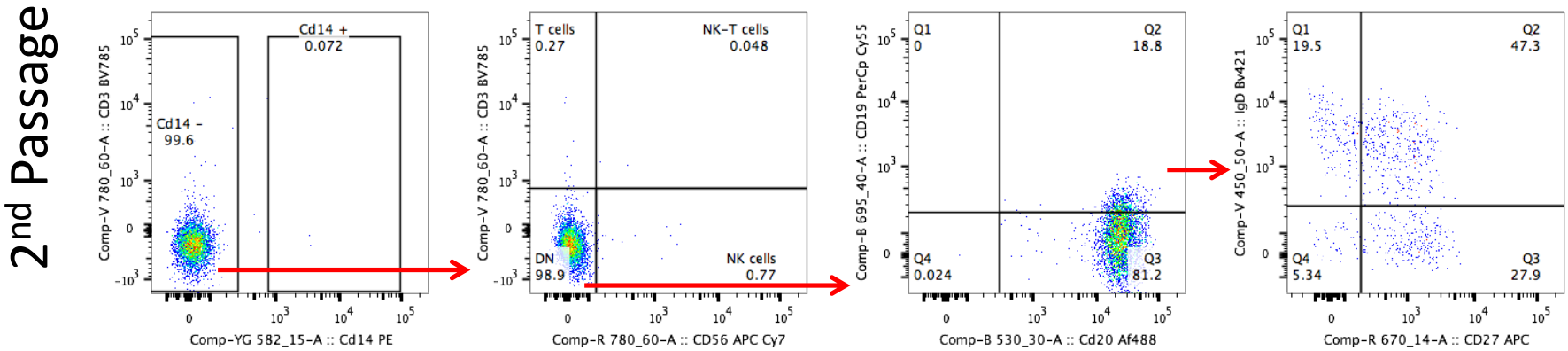


Figure 4.9. 3nd FACS Panel, double negative population issue solved (previous pages, first page PBMCs extraction, second page 1st separation, third page 2nd separation)): The inclusion of CD20 in my 3rd FACs panel have allowed me to discriminate the CD19+ before and after the purification, once the CD19+ epitope was masked by the purification antibody. Cells have been gated on physical parameters, then singlet, Aqua-, CD14-, CD3-/CD56-, CD20+ or CD20+/CD19+.

Final panels: Purification, Characterization and Lymphoblast panel

At this point, in order to standardise my co-culture system, I decided to setup 3 final panels that would have help to assess purity, but also to characterize my B cells subsets at the beginning and the end of the co-culture.
These 3 panels are: i) Purity Panel (basically the same of my 3rd panel); ii) Characterization panel (this panel is the most important, containing not just CD38, IgD and CD27, but also IgG , IgM and CD24); iii) Lymphoblast panel (this panel contain normal B cells marker plus the addiction of CD23 and CD58, specific for EBV relates subsets as already mentioned in the introduction (page 91), whose outcome will be discussed in chapter 6. All the 3 panels can be visualized in appendix in table 3.2.

Optimization of FACS on FLS

Increased number of FLS and surface marker preservation for FACS

I decided to focus also on the FLS and to setup a working FACS panel with the aim of exploring the effect induced by the B cells to the FLS during the time in culture. In order to optimize this procedure, I used fibroblast from different sources (RA, OA, PsA). The markers selected for this panel are both the more typical FLS markers (CD90, Cadherin 11), plus some of the ones that have gained interest recently (Podoplanin, DAF/CD55 and FAP). I also included Glycophorin (GYPA), CD45, PECAM1 and CD146 in the dump channel.
After performing numerous attempts of optimization, the first obstacle has resulted in obtaining enough cells for a correct analysis. Indeed, when starting from a number of FLS comparable to PBMCs, the former had always resulted in a much lower yield. Luckily, according to Kaur and Esau report, I discover that almost 80% of stromal cells are lost during a

normal staining procedure, due to centrifugation [250]. Following the guidelines from this paper, I reduced the number of washes and centrifugation, and this had led to a strong increase in the fibroblast yield (Fig 4.10).

In the same context I had to solve another issue. Compared to PBMCs, FLS are adherent cells and thus require being detached from the plate. Unfortunately, the use of trypsin has resulted in an almost total loss of some important surface marker, such as CD90. In order to solve this problem, I tested several detaching agents (both enzymatic and non-enzymatic) such as accutase or versene, with the latter showing the better results in term of signal preservation for all the markers (Fig 4.11). It is important to mention that - compared to trypsine and accutase - versene require longer incubation times and this might lead to and increase number of dead cells, due to toxicity. Reduction of these incubation times and adding PBS 1:1 to the cell-versene suspension has increased viability. Interestingly, the fact that versene is a gentler detaching agent has resulted of great help for my co-culture system, in which I wanted to analyse the B cells and the FLS separately. Since B cells after long time and culture tend to be attached to the stromal cells it would have been difficult to sort them using trypsin. The use of versene and the longer incubation time has allowed me to extract from the co-culture first an enriched B cells fraction (with few FLS) and then an enriched FLS fraction (with a mild B cells contamination).

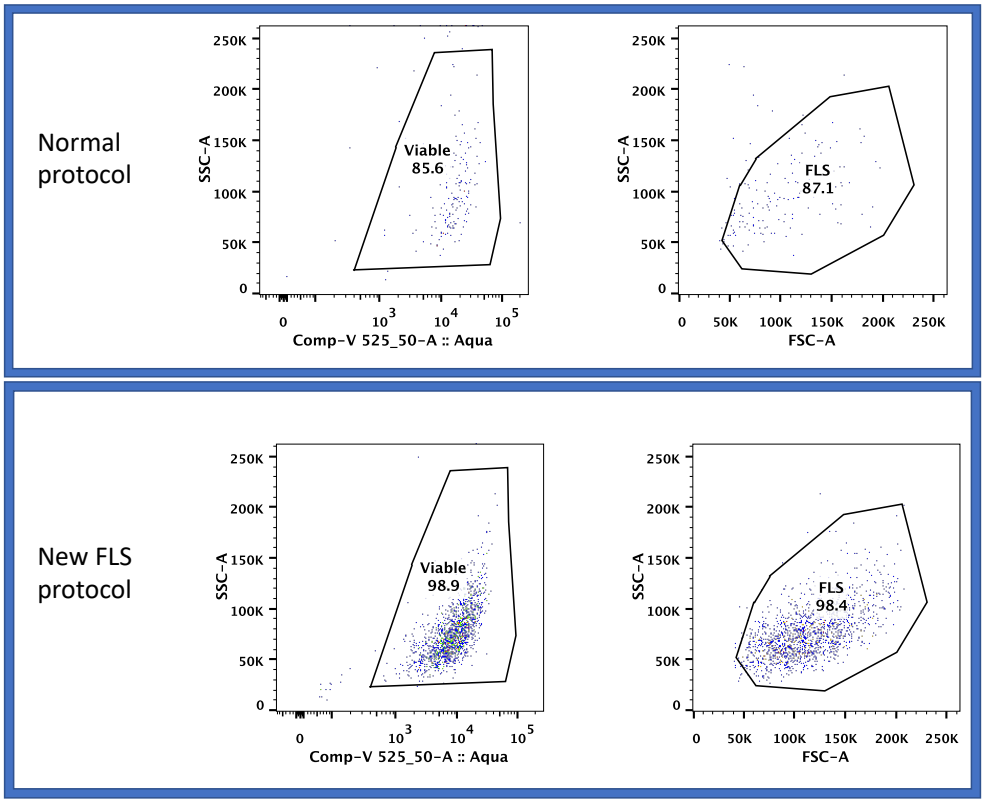
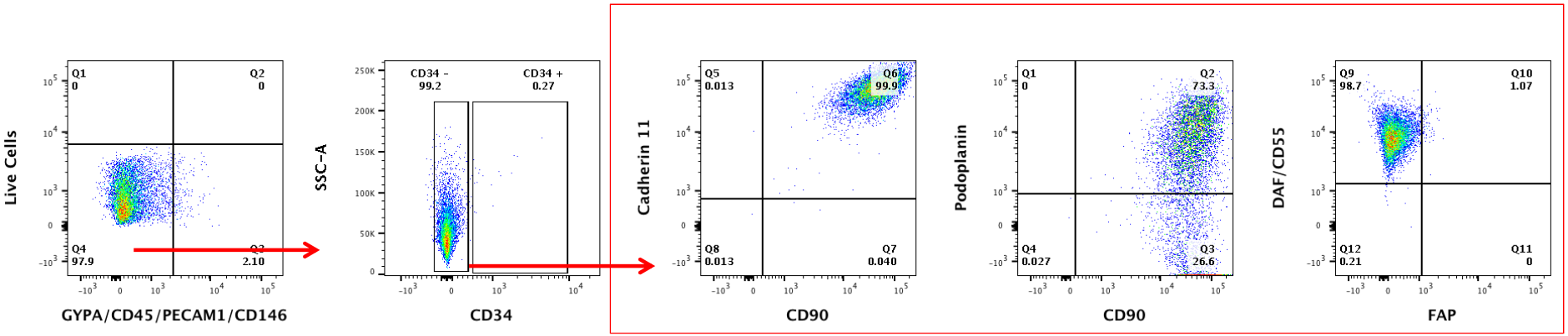
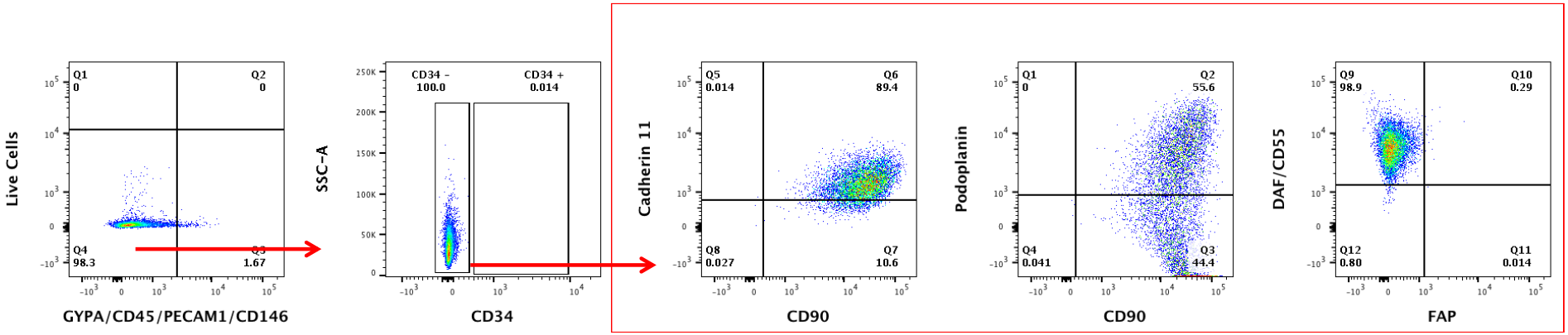


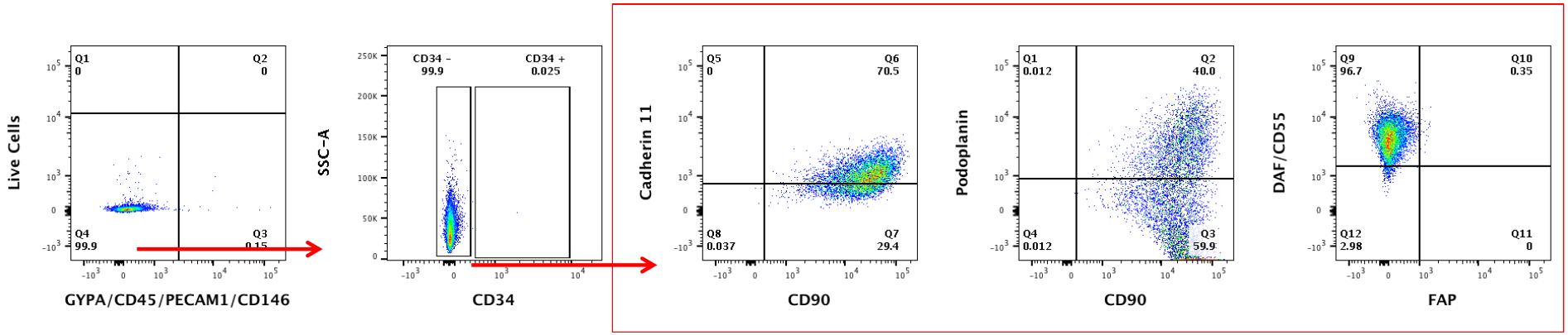
Figure 4.10. Increased number of cells with New FLS Protocol: After reducing the number of centrifugation steps in the original protocol I managed to obtain a correct number of cells. Cells are gated for physical parameters, then for Zombie aqua.


Live Cells
GYPA/CD45/PECAM1/CD146
Q1 0
Q2 0
Q4 97.9
Q3 2.10
SSC-A
CD34
CD34 - 99.2
CD34 + 0.27
Cadherin 11
CD90
Q5 0.013
Q6 99.9
Q8 0.013
Q7 0.040
Podoplanin
CD90
Q1 0
Q2 73.3
Q4 0.027
Q3 26.6
DAF/CD55
FAP
Q9 98.7
Q10 1.07
Q12 0.21
Q11 0

Versene


Live Cells
GYPA/CD45/PECAM1/CD146
Q1 0
Q2 0
Q4 98.3
Q3 1.67
SSC-A
CD34
CD34 - 100.0
CD34 + 0.014
Cadherin 11
CD90
Q5 0.014
Q6 89.4
Q8 0.027
Q7 10.6
Podoplanin
CD90
Q1 0
Q2 55.6
Q4 0.041
Q3 44.4
DAF/CD55
FAP
Q9 98.9
Q10 0.29
Q12 0.80
Q11 0.014

Accutase



Trypsin

Figure 4.11. FLS detaching protocol: comparison of different detaching agents (previous pages; first page: versene separation, second page: accutase separation; third pages: trypsin separation) a): the same RA-FLS sample have been detached with versene (V), accutase (A) and trypsine (T). 10'000 events have been analysed. Cells have been gated on physical parameters, Zombie Nir/dump channel, CD34-, Cadherin11/CD90, Podoplanin/CD90, DAF/FAP.

Long-term culture of FLS cause small phenotype variations and difficult reproducibility of FACS data

Once the right protocol was set-up, I decided to test both the co-culture system FLS, and to compare FLS from a different autoimmune disease to investigate any variability in my set of markers. In my first experiment I compared FLS from PsA and RA and the first result has been encouraging (Fig 4.12). Despite being identical for almost all the other markers (Cadherin11, Podoplanin, CD34 and CD90, the RA-FLS showed a clear double positive population for DAF/CD55 and FAP when compared to PsA-FLS. I decided to do the same for OA-FLS and also to test if the same samples can be stable or show differences trough passages. Unfortunately, this has not been the case: Figure 4.13 is showing that the same sample of RA-FLS or OA-FLS behave quite randomly regarding the expression of this surface markers, with small differences between the two disease.
Also, most of the small differences obtained - both when comparing cells from different diseases and passages or in my co-culture- have been extremely difficult to reproduce and are mostly more related to the time in culture and to experiment-induced effects.

The co-culture results concerning the sorted RA-FLS at the end of each experiment have followed a similar pattern and further analysis is needed.

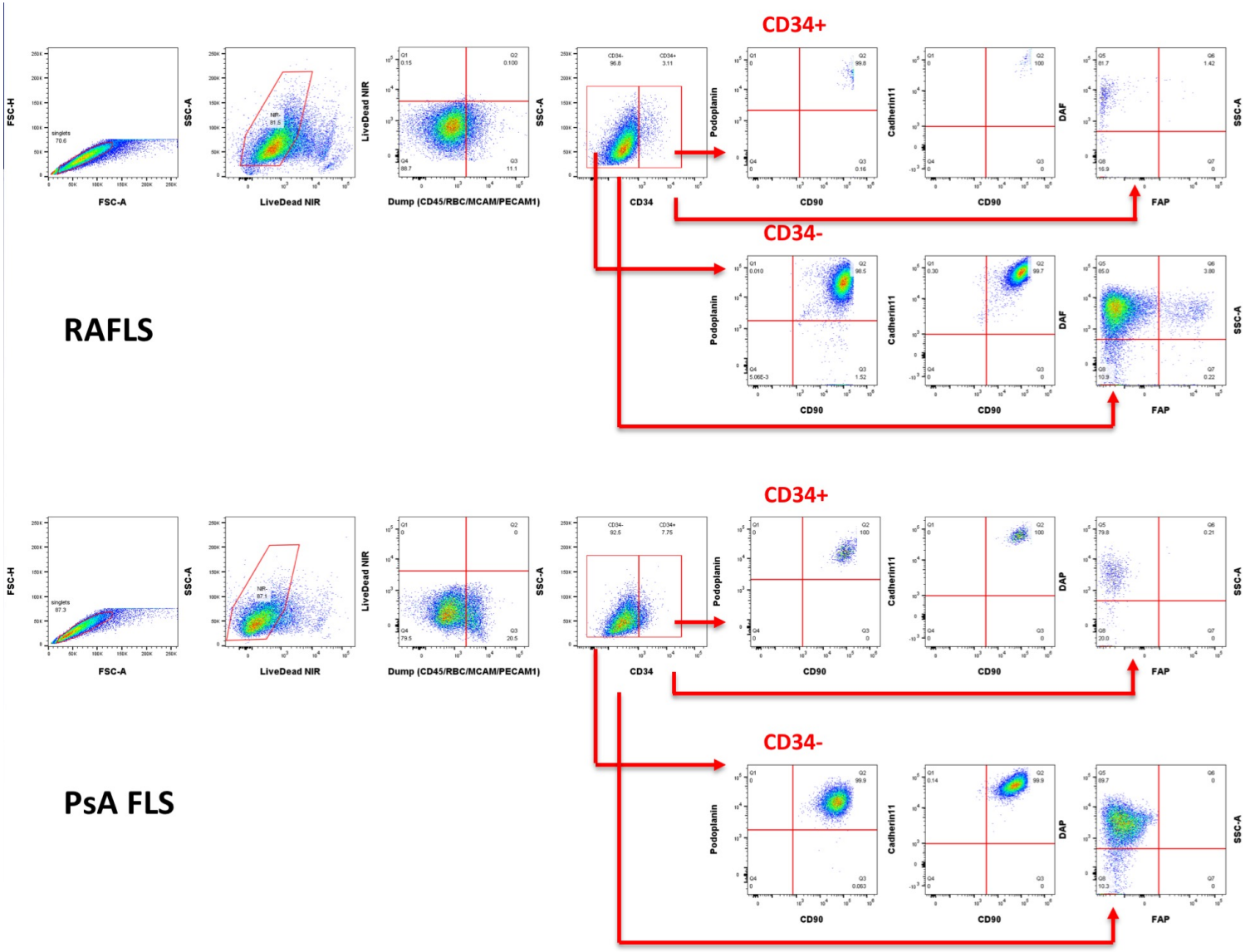
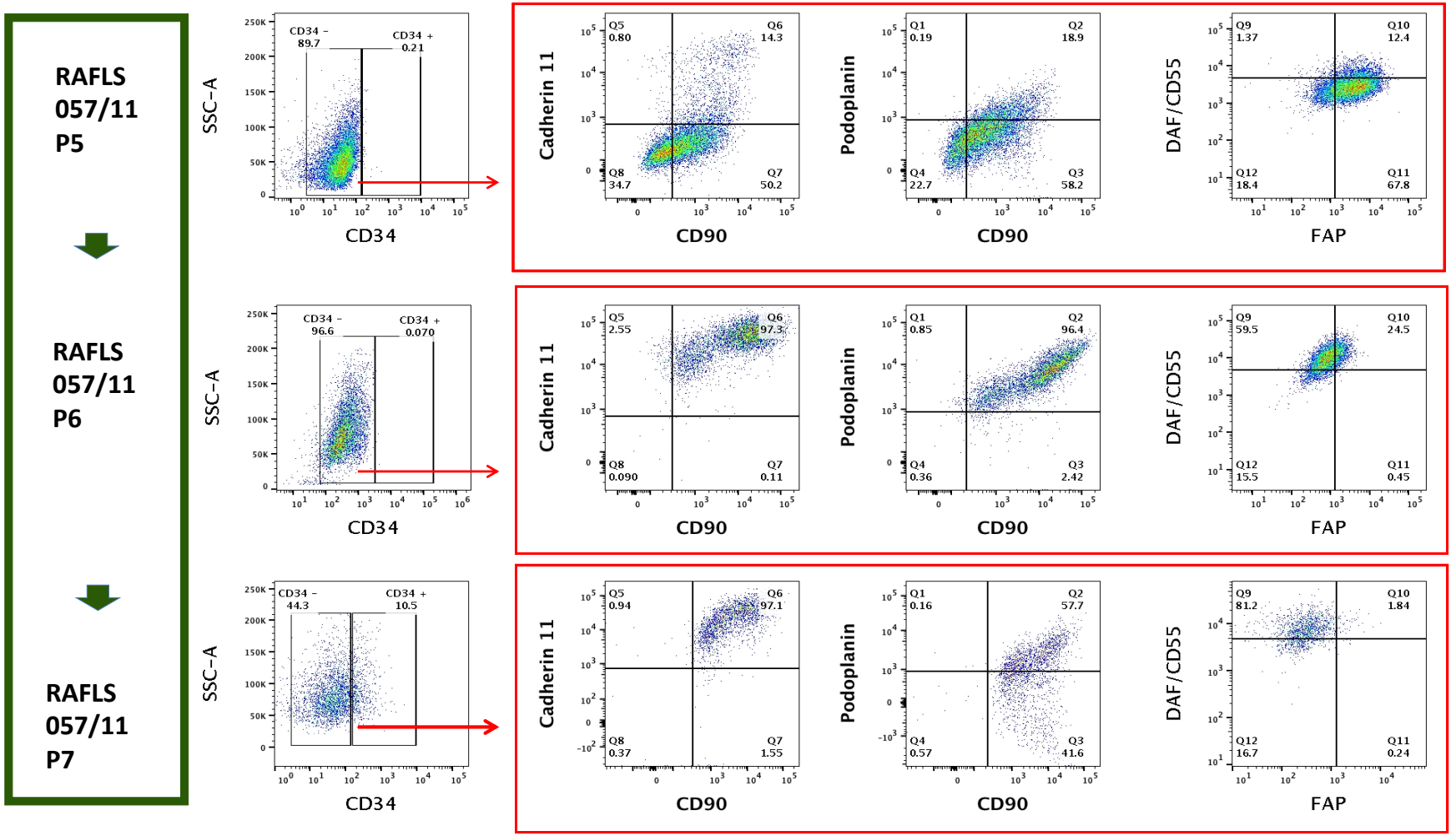

CD34+
CD34-
RAFLS
PsA FLS
FSC-H
FSC-A
SSC-A
LiveDead NIR
Dump (CD45/RBC/MCAM/PECAM1)
CD34
Podoplanin
Cadherin11
CD90
DAF
DAP
FAP

Figure 4.12. FACS Comparison between RA-FLS and PsA-FLS, passage 5 (previous page): RA-FLS and PsA-FLS have been compared trough FACS. The two samples appear quite similar regarding the expression of surface markers CD34. CD90, Podoplanin and Cadherin 11. Nevertheless RA-FLS show a double population of DAF/FAP which is absent PsA-FLS. Cells have been gated on physical parameters, Zombie Nir/dump channel, CD34-, Cadherin11/CD90, Podoplanin/CD90, DAF/FAP.


RAFLS 057/11 P5
RAFLS 057/11 P6
RAFLS 057/11 P7
SSC-A
CD34
CD34 - 89.7
CD34 + 0.21
CD34 - 96.6
CD34 + 0.070
CD34 - 44.3
CD34 + 10.5
Cadherin 11
Podoplanin
DAF/CD55
CD90
FAP
Q5 0.80
Q6 14.3
Q8 34.7
Q7 50.2
Q1 0.19
Q2 18.9
Q4 22.7
Q3 58.2
Q9 1.37
Q10 12.4
Q12 18.4
Q11 67.8
Q5 2.55
Q6 97.3
Q8 0.090
Q7 0.11
Q1 0.85
Q2 96.4
Q4 0.36
Q3 2.42
Q9 59.5
Q10 24.5
Q12 15.5
Q11 0.45
Q5 0.94
Q6 97.1
Q8 0.37
Q7 1.55
Q1 0.16
Q2 57.7
Q4 0.57
Q3 41.6
Q9 81.2
Q10 1.84
Q12 16.7
Q11 0.24

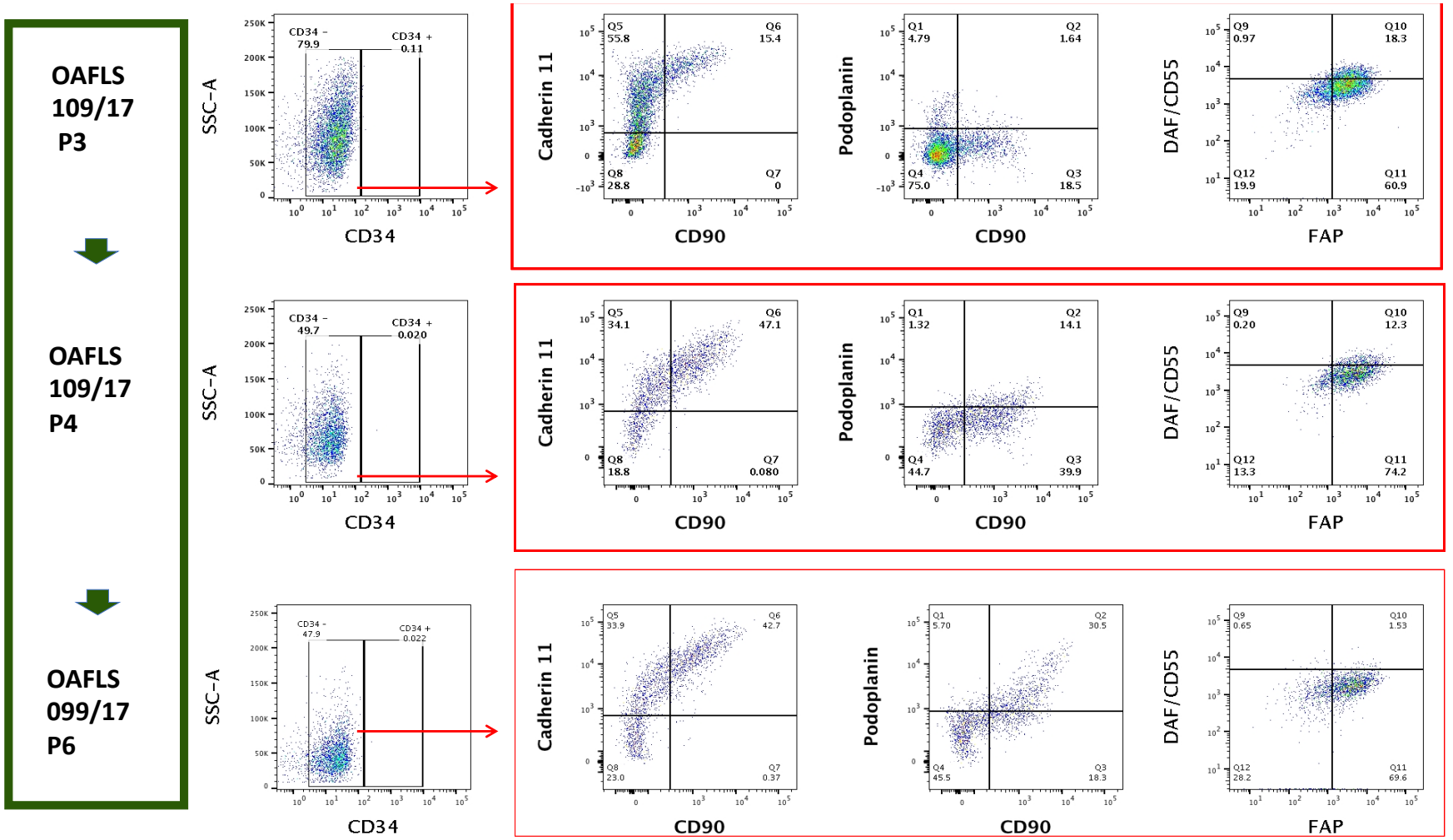

OAFLS 109/17 P3
OAFLS 109/17 P4
OAFLS 099/17 P6
SSC-A
CD34
CD34 - 79.9
CD34 + 0.11
CD34 - 49.7
CD34 + 0.020
CD34 - 47.9
CD34 + 0.022
Cadherin 11
CD90
Q5 55.8
Q6 15.4
Q8 28.8
Q7 0
Q5 34.1
Q6 47.1
Q8 18.8
Q7 0.080
Q5 33.9
Q6 42.7
Q8 23.0
Q7 0.37
Podoplanin
Q1 4.79
Q2 1.64
Q4 75.0
Q3 18.5
Q1 1.32
Q2 14.1
Q4 44.7
Q3 39.9
Q1 5.70
Q2 30.5
Q4 45.5
Q3 18.3
DAF/CD55
FAP
Q9 0.97
Q10 18.3
Q12 19.9
Q11 60.9
Q9 0.20
Q10 12.3
Q12 13.3
Q11 74.2
Q9 0.65
Q10 1.53
Q12 28.2
Q11 69.6

Figure 4.13. FACS Comparison between RA-FLS and OA-FLS at different passages (previous pages; first page: RA-FLS; second page: OAFLS): RA-FLS 057/11 and OA-FLS 109/17 and 099/17 have been compared trough FACS at different passages. The small differences observed - regarding the expression of surface markers CD34, CD90, Podoplanin and Cadherin 11 - between the two batches of sample appear to be random and mostly more related to the time in culture and to experiment-induced effects. Cells have been gated on physical parameters, Zombie Nir/dump channel, CD34-, Cadherin11/CD90, Podoplanin/CD90, DAF/FAP.

FLS screening to address B cell-targeting stimuli induced effect in the co-culture

CpG and IL-2 do not influence neither expression nor inflammatory cytokines production in RA-FLS

During preliminary experiments for the co-culture system, a small difference has been observed in term of IL-6 response between un-stimulated and stimulated RA-FLS with CpG and Il-2. This result was puzzling since FLS are usually thought to lack both IL-2 receptor and TLR9 [118]. In order to address this, a test has been run: RA-FLS have been cultured for 28 days in the same exact conditions obtained during co-culture, with or without the stimuli. SN was collected every 7 days and - at the end of experiment - cells have been digested to obtain RNA. The SN has been tested at the different time points for the expression of IL-6 or IL-8. The cDNA has been tested on qPCR for expression of IL-6 and IL-8, but also for TLR-9, BAFF, APRIL and IL2 receptor subunits (IL-2Rβ and ILR-2γ).

The results showed - both by ELISA and by qPCR – that the stimuli do not induce IL-6 or IL-8, so the small difference observed was probably a random effect (Fig.4.14 & 4.15). Also, RA-FLS seemed to have very basal levels of TLR-9 and the IL-2R subunits, which are not altered by CpG or IL-2.

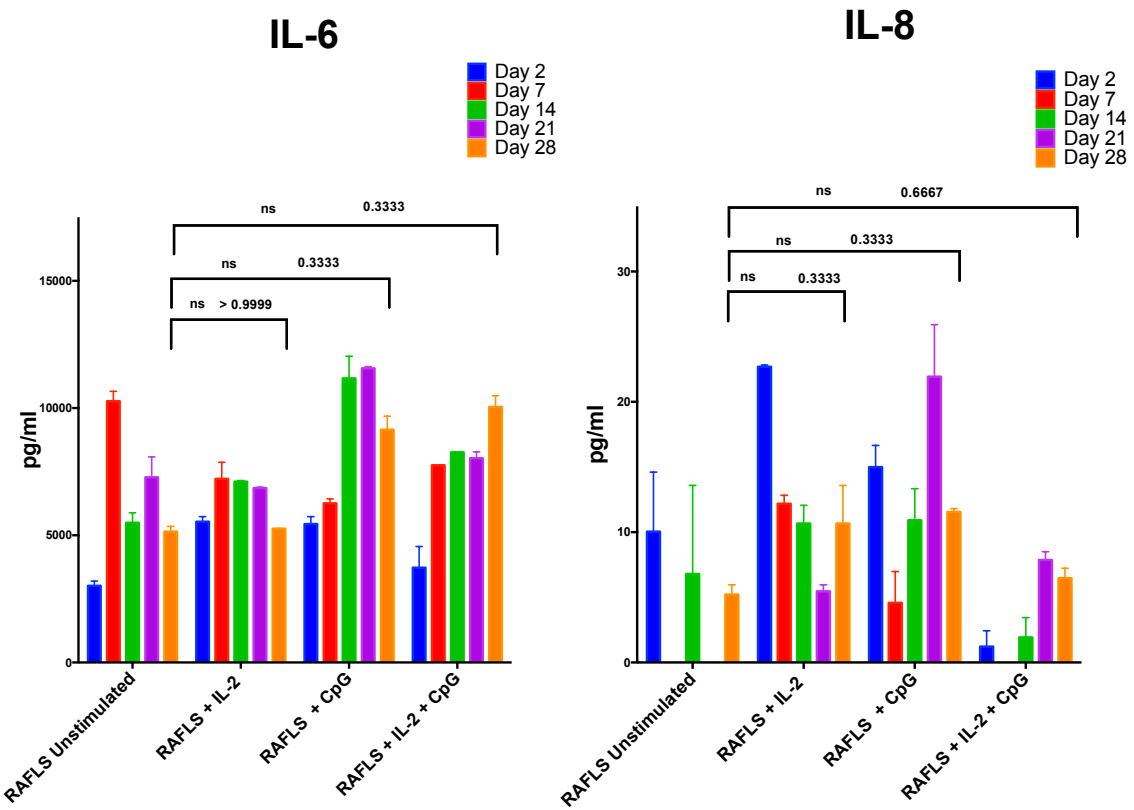


Figure 4.14. IL-6 and IL-8 production in RA-FLS stimulated with CpG and IL-2: Results are showing that CpG and IL-2 have no effect on RA-FLS.

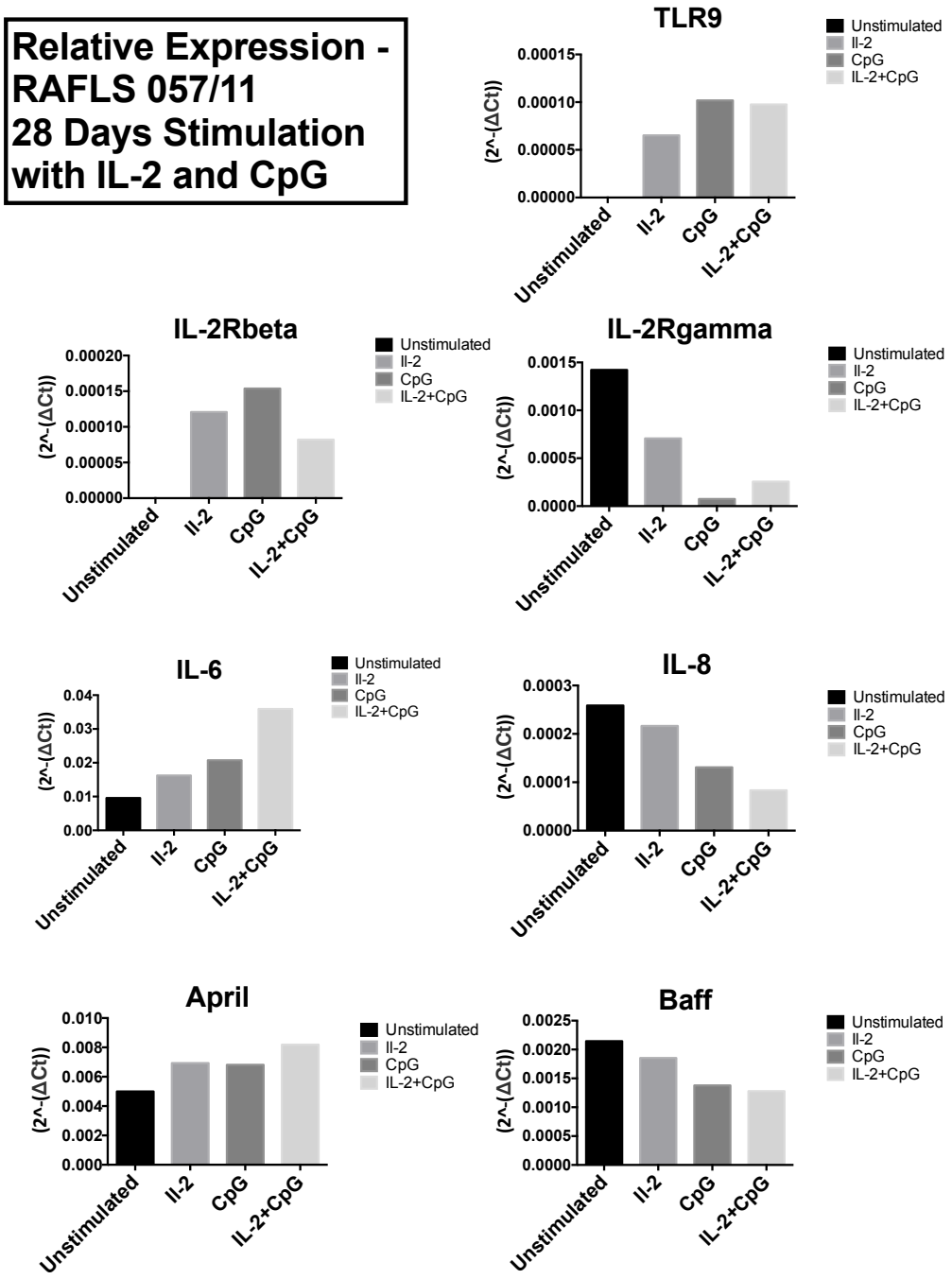


Figure 4.15. Relative Expression of co-culture related genes after 28 days stimulation with IL-2 and CpG on RA-FLS: Results relative expression on TLR9, IL-2Rβ, Il2-Rγ, BAFF, APRIL, IL-6 and IL-8.

First proliferation events obtained in co-culture system

Proliferation of RACD19+ B-cells is obtained in the presence of RA-FLS

After 40 days of culture, in two experiments out of six, a sudden increased in the number of B cells was observed. This particular behaviour was observed only in 1 and 3 of the 8 biological replicates (Patient 1 3/8; Patient 6 1/8) (Fig 4.16)

The confirmation that the phenomenon observed was actually proliferation - and not only activation - has been given from the number of events obtained during FACS analysis and single cells sorting at the end of the experiment: considering a starting number of 12'500 RACD19+ cells (p/w), at least in one well analysed 65'192 events were reached, giving a 5.21 ratio (the latter is not actually taking into account any death ratio, which are obviously believed to take place during this long experiment).

Another important improvement at this point has been the introduction of the "Time 0" control: at the beginning of every co-culture, a small aliquot of isolate CD19+ B cells has been frozen to be tested as a control both for B cells subsets variation and for the presence of EBV. Unfortunately, at the time, none of my assays was working properly to identify and quantify the virus. This was later resolved and data comparing baseline and post-co-culture EBV content are presented in Chapter 5.

Nevertheless, results have shown that the majority of proliferating B-cells where alive and CD38+, when compared to the Time 0 Ctrl cells (Fig 4.17) This marker is usually expressed by proliferating lymphoblast in the GC but is also highly express by lymphoblastoid cell line which are established in vitro from the survival and proliferation of B cell naturally harbouring the EBV [251].

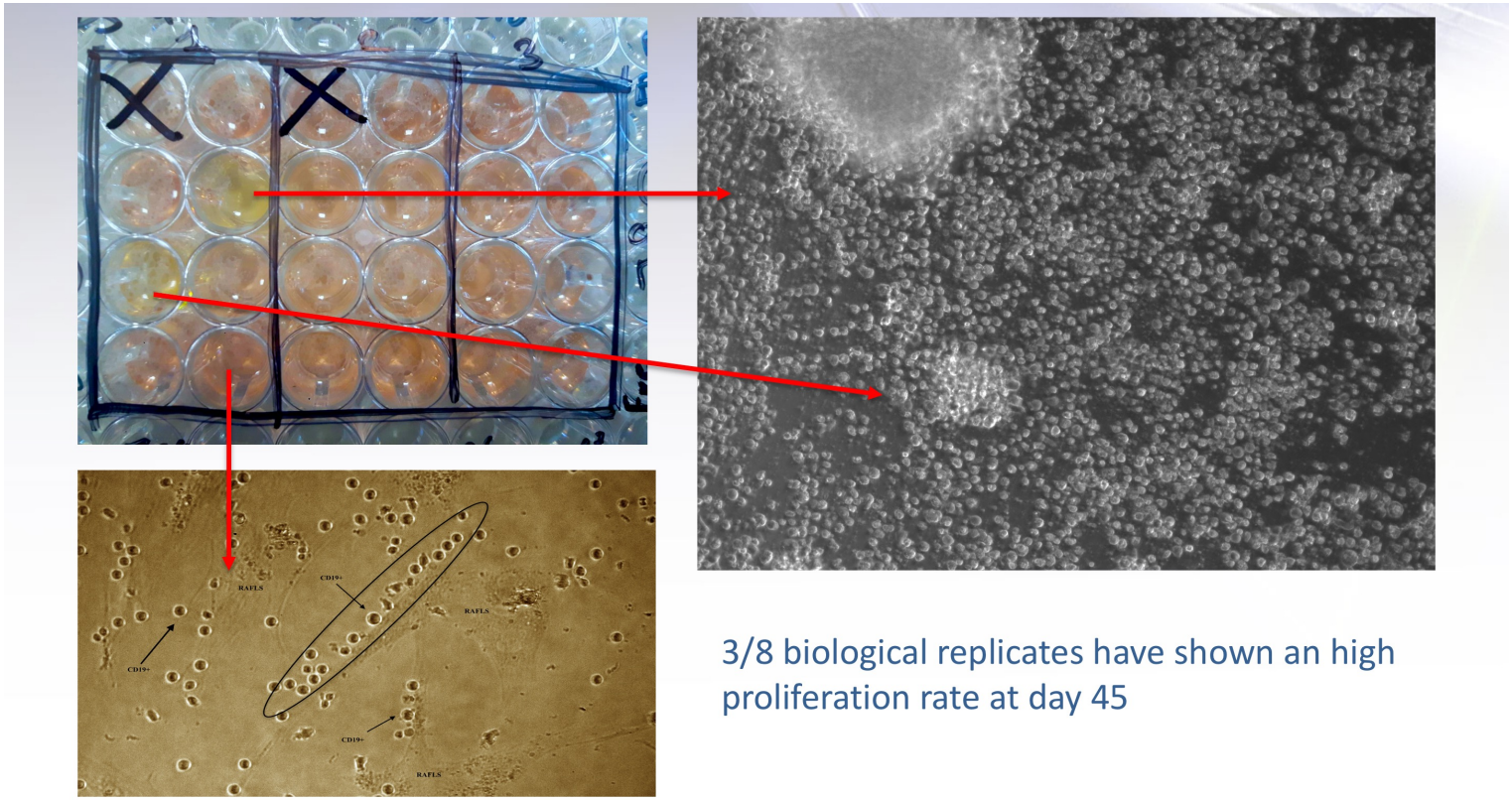
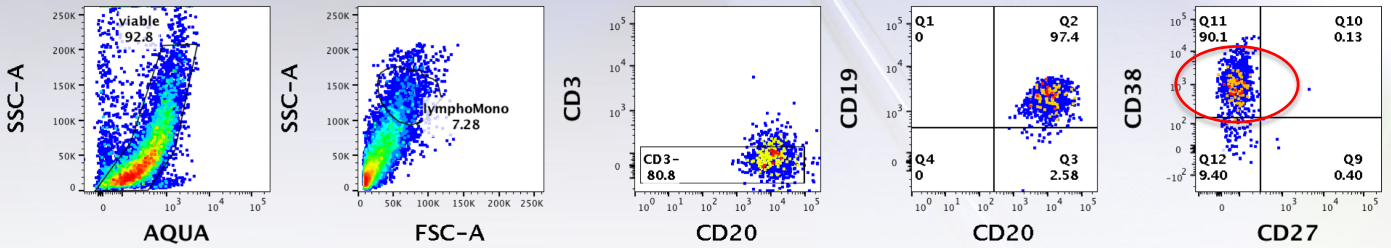


Figure 4.16. Proliferation of CD19+/CD38+ in 3wells on day 45. 3 out of 8 biological replicates have shown a high proliferation rate at day 45.

RACD19+
(co-cultured with RAFLS)



- Presence of live cells
- CD19+ /CD20+
- Increase in CD38+ Lymphoblastoid Cells

RACD19+
(Ctrl - Time 0)

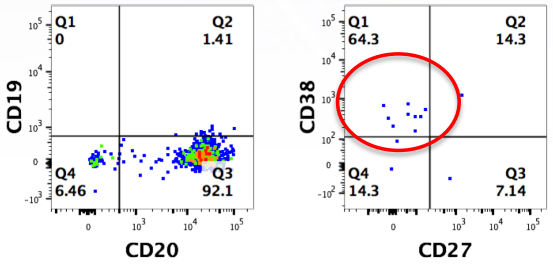


Figure 4.17. Sorting of Lymphoblastoid CD38+ from co-culture (previous page): Summary slide of final sorting on CD19+ B cells obtained at the end of the co-culture (day 105) compared to the CD19+ B cells purified at time 0. Cells are alive and the EBV-driven proliferation has resulted in a lymphoblast/GC CD38high alike cells. Gating strategy: cells have been gated on aqua for viability; on side scatter (SSC-A) and forward scatter (FSC-A) for physical parameter. Then CD3-CD20+ to exclude T cell contamination. Finally, on CD19+CD20+ and on CD38+CD27- for lymphoblast/GC CD38high alike cells. Interestingly at time 0 cells are CD19-CD20+ and this is direct consequence of the anti-CD19 antibody (milteny) - used for positive selection at the beginning of the experiment – masking the CD19 on the cells surface.

CHAPTER 5: RA-FLS support proliferation, activation and differentiation of naturally occurring EBV+B cells

In order to identify and quantify EBV in my co-culture system, several attempts have been performed with different molecular biology protocol. Unfortunately, most of them have been unsuccessful until finding the most suitable one. In these attempts I have i) tested the "old" batch of primer mentioned in chapter 4; ii) create a protocol for a gradient PCR; iii) create a protocol for a Nested PCR. None of these tests have worked due to lack of signal or specifity (data not shown).

A reliable method for EBV quantification: BAMHI w repeated region sequence gDNA qPCR

Successful Preliminary test for BAMHI w repeated sequence gDNA qPCR and sequencing results

Thanks to the kindness of Dr Thorley Lawson, who has helped me in solving this issue, I later found out that the actual state of the art for EBV analysis and quantification it is a specific genomic DNA qPCR targeting the BAMHI W repeated region and based on limiting dilutions technique [243].

Once obtained the probe and the cells to be used as control (new batches of Namalwa cells and Ramos), I decided to test the co-culture in which I observed proliferation. Several aliquots of frozen cells from these experiments have been thawed and then compared to a standard curve obtained from limiting dilutions of Namalwa cells. Indeed, this further

molecular analysis demonstrated that all sorted proliferating cells were positive for EBV (Fig 5.1).

Unfortunately - since these cells were bulk sorted - a precise interpolation of the number of infected cells at a single cells level was not possible. Nevertheless, the qPCR template has then been used to perform electrophoresis analysis, which allowed the identification of a clear 191 bp product, absent in the negative control (Fig.5.2). This band was then sequenced, giving the final confirmation, by an alignment of 100 %, that the product matched with the Epstein Barr Virus genome (Fig 5.3).

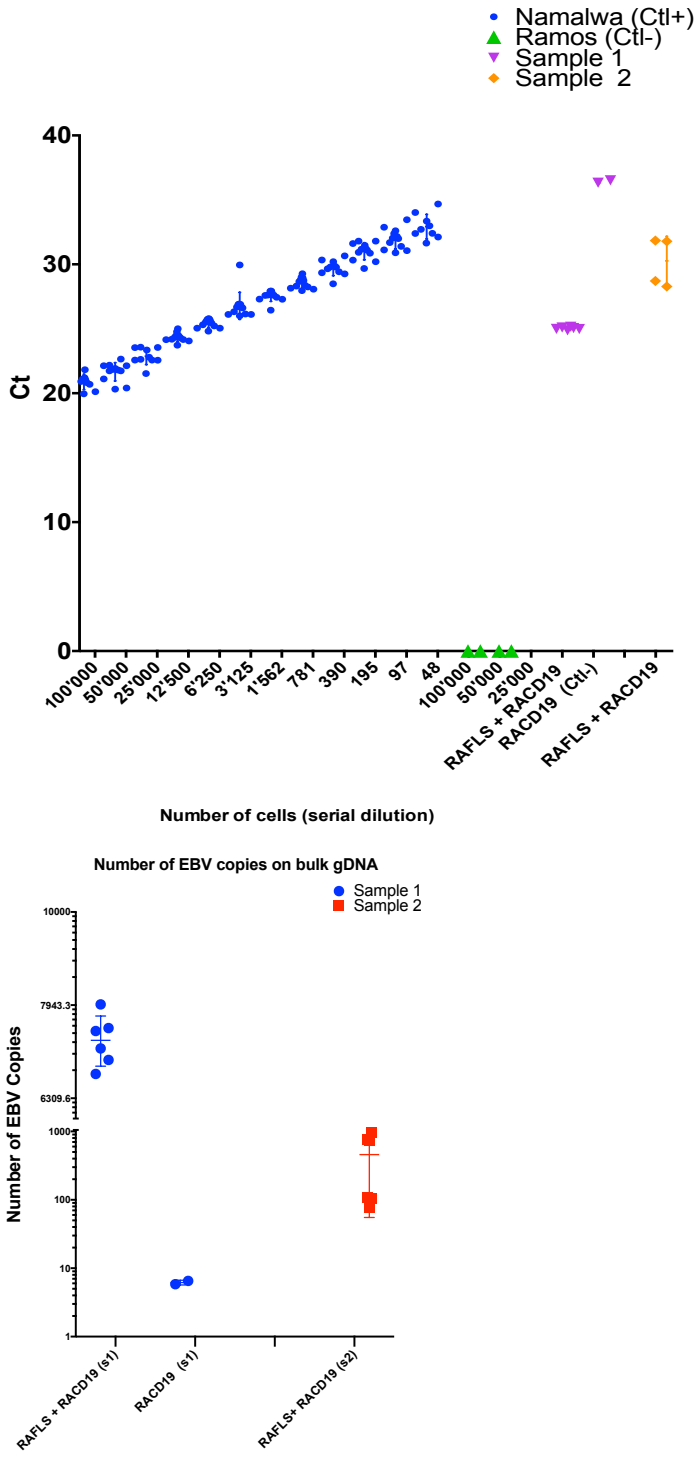


Figure 5.1. PCR for the BamH1 W repeat region of EBV (W-PCR) and calculation of frequency/level of EBV copies in bulk gDNA (Thorley-Lawson Laboratory). Cells sorted from proliferative wells resulted to be all positive for EBV. Also, number of infected cells was higher when compared to internal control. Namalwa cells and Ramos cell were used as a positive and negative control for EBV. Limitations of this assay is the absence of a Time 0 control and the fact that I had bulk gDNA instead of precise cells number.

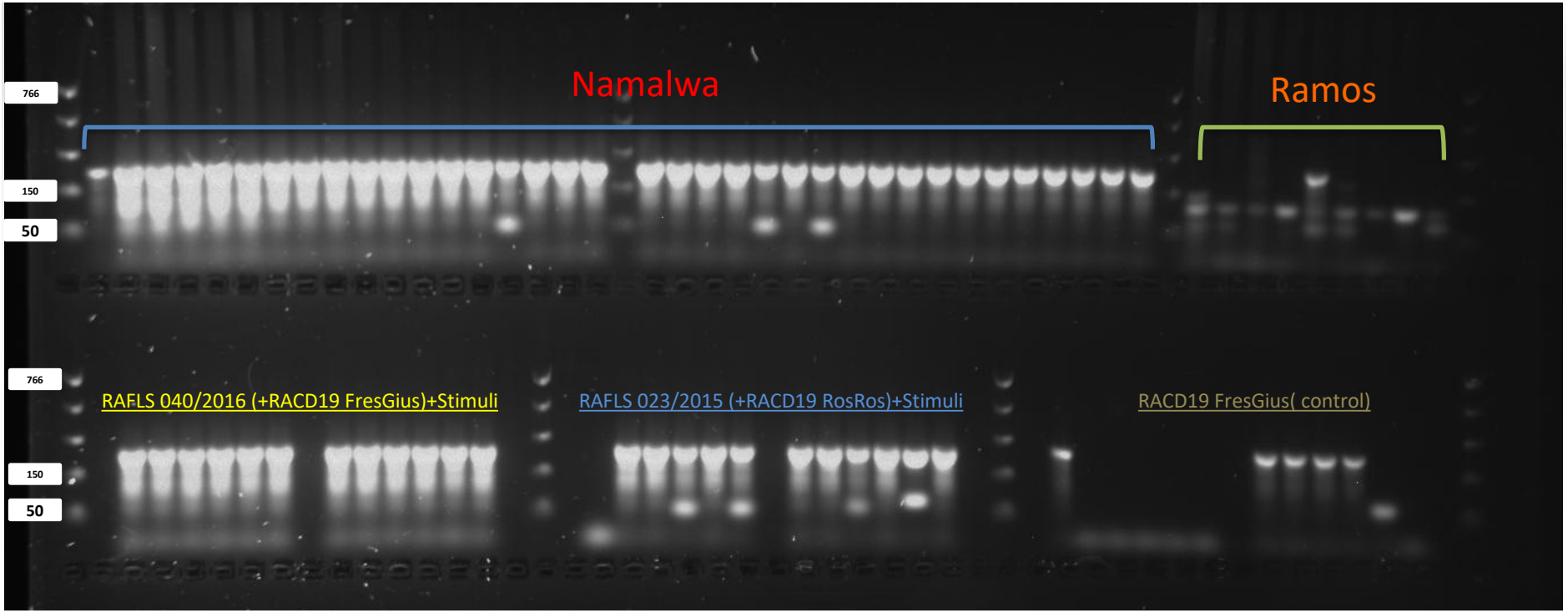


Figure 5.2. Agarose gel on W-PCR products: In positive control the product of the qPCR clearly appears at 191bp. This was absent in the negative control but present in 2 patients samples. The fact that this procedure is genomic a DNAqPCR explain the high presence of background. Occasionally the qPCR probe can appear as a second product between 50bp and 20bp.

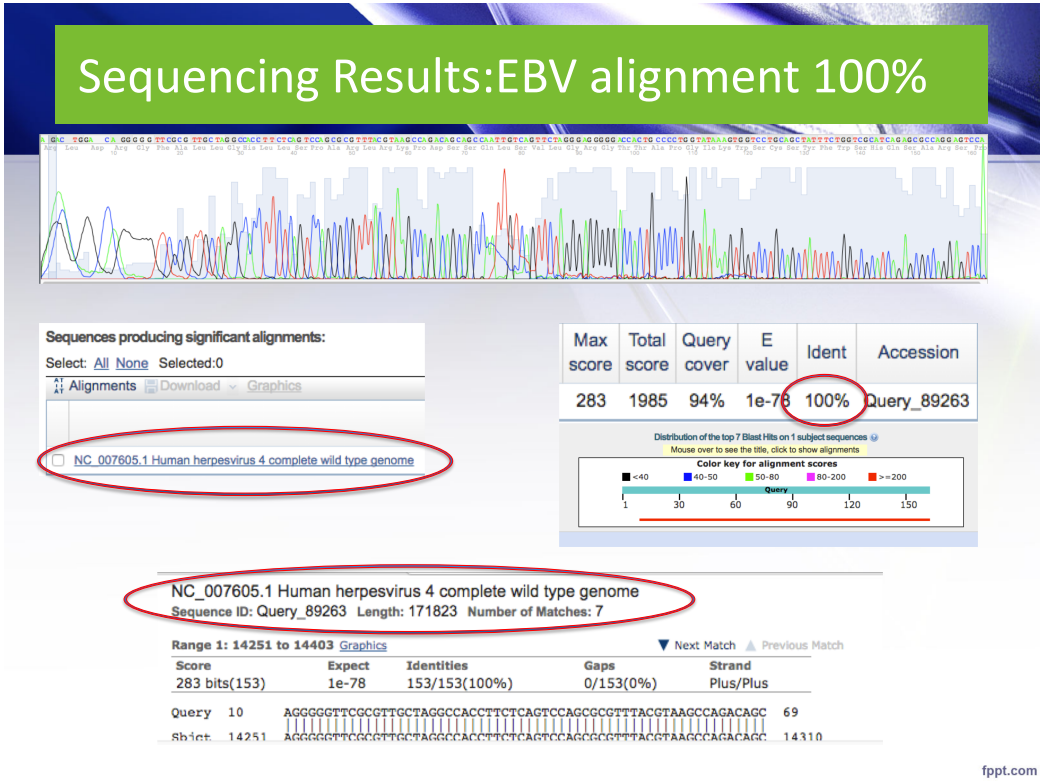


Figure 5.3. Sequencing on W-PCR products: Slide summary for sequencing results on W-PCR products. The 191bp product has been used to perform sequencing analysis giving an alignment of 100% with Human Herpesvirus 4 complete wild type genome.

Optimization test on Namalwa cells to improve EBV BamHI gDNA-qPCR at single cell level

In order to improve the BamHI qPCR at a single cells level I decided to test whether different combination of gDNA and water can alter the outcome of my reactions. Also, I decided to sort on live cells for the standard curve, instead of using only the physical parameters. The standard curve of Namalwa cells has been performed manually starting from 50'000 cells and performing serial dilution till 19 cells. For 7, 2 and 1, the cells have been sorted. With a 10 µl final volume, the best results have been obtained when using water, a lower amount of template and when cells were sorted starting from the live ones (or Zombie aqua negative) (Fig 5.4).

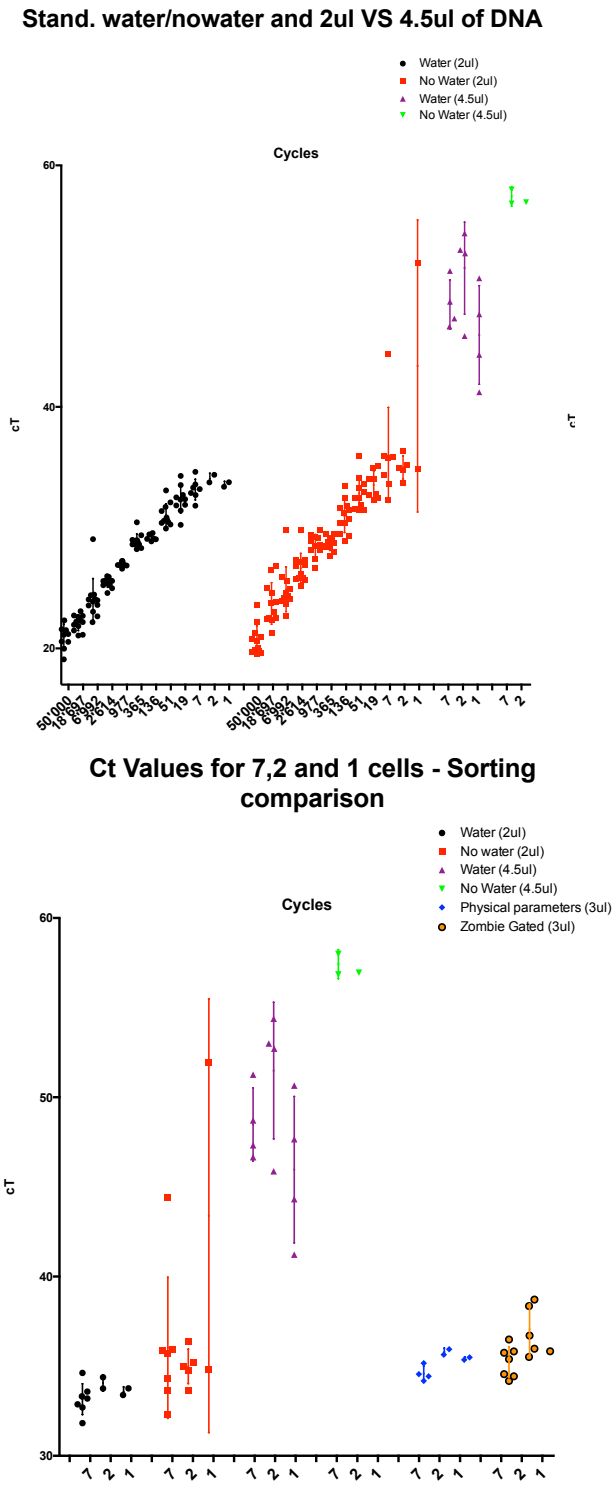


Figure 5.4. Optimization test on Namalwa cells to improve EBV BamHI DNA-qPCR at single cell level: Comparison between limiting dilutions of Namalwa cells screened with different reaction mix. For conditions "4.5ul" with or without water and for "zombie gated" (live cells) VS "physical parameters" only 7,2 and 1 sorted cells have been used.

Experiment 0051: Increased Number of CD19+ and first evidence that proliferation of CD19+ B-cells is EBV driven and obtained in the presence of RA-FLS:

At this point I needed to standardise my experiment and repeat it. As mentioned in research methodologies, I also modified some of the conditions, mostly to increase the probability to have EBV+CD19+ in my co-culture.

Indeed, my first test with the new conditions, experiment 0051, has shown proliferation after only 14 days. Sorting has been performed at day 28 and cells have been tested with BamHIW EBVqPCR. Thanks to the presence of a control at time 0, it has been possible to check the increase in number of EBV copies compared to the beginning of the experiment, and also to esteem the average of copies for single cells (Fig 5.5). Limiting dilutions have been performed both for samples and controls, in order to identify the lowest point in which was possible to still obtain a positive signal.
Results clearly show an increase in number of EBV copies compare to control (200'000 copies at 28 days Vs 10-1 copies at time 0) using the same number of cells. A low level of proliferation has been observed also in the B cell control but this data has to be revised since no alive B cells where present in this condition and most of the cells sorted were dead or dying. It is very likely that by adding DNAse to the cells - before performing the sorting - this kind of background can be avoided. Data have then been confirmed on an agarose gel (data not shown).

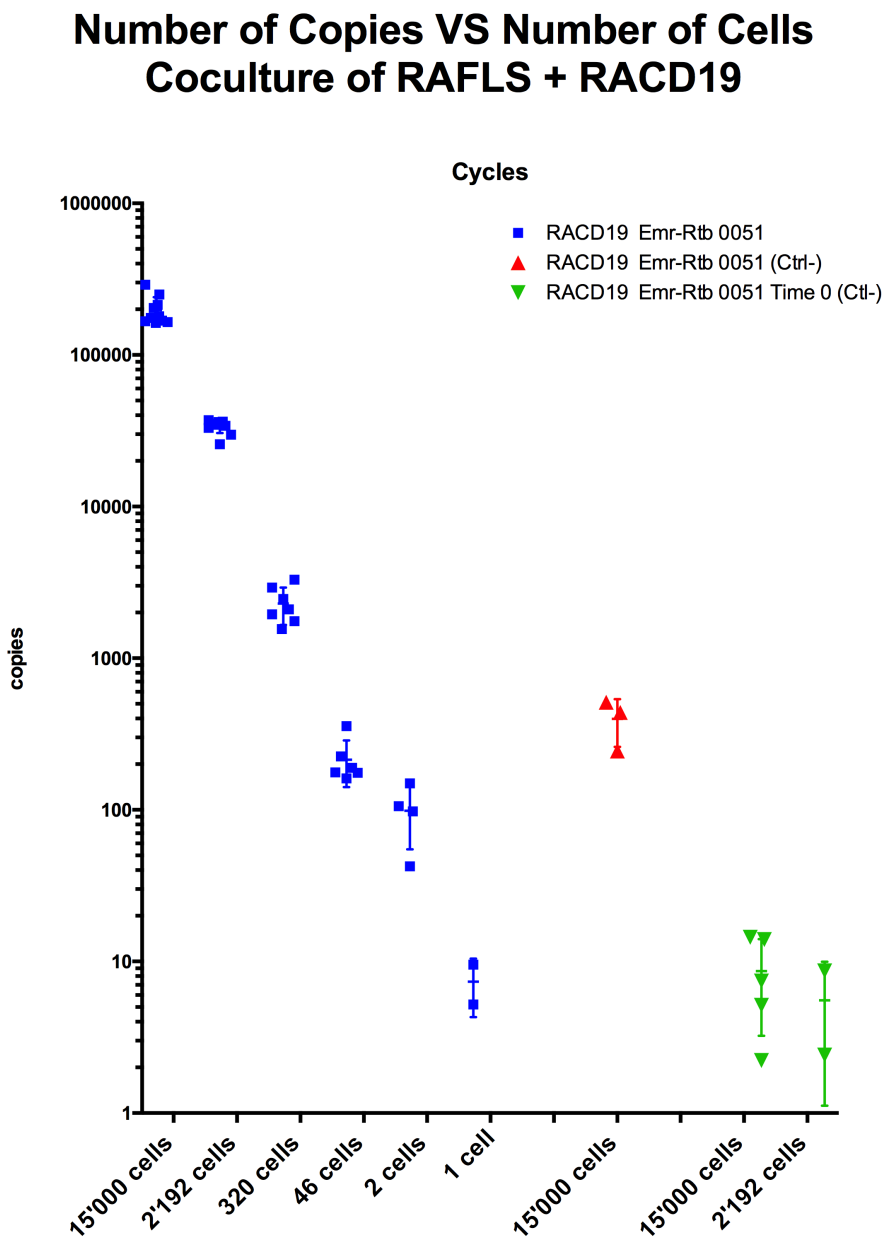


Figure 5.5. Number of Copies versus Number of B cells in co-culture system: Results are showing an enormous increase of the infected cells and number of copies after co-culture with RA-FLS compared to control. Each sample has been run in 4 biological replicates and each of the biological replicates has been further tested in 3 technical replicates for BAMHIW EBVqPCR. Symbols represent technical replicates

Comparison between co-culture of RA-FLS with HdCD19+ or RACD19+

Co-culture of RA-FLS with RACD19+ and HdCD19+ B cells shows increase of EBV infected cells

In order to address one of the aims of this PhD I repeated the experiment 10 times using 5 blood samples from ACPA+ RA patient and compared these results VS RA-FLS cultured with HdCD19+ obtained from 5 different healthy donor. Each of these 10 co-cultures have been performed separately and not at the same time, as previously stated. After the last experiment have been performed data have been put together to create this graph. I confirmed that after 28 days, proliferation of B cells happens exclusively when cultured with RA-FLS and it is always due to the presence of EBV. I also demonstrated that this EBV-driven proliferation does not happen exclusively with RACD19+ but also with HdCD19+ (as this was expected since circulating EBV+ B cells are present at a very low frequency also in this group). Nevertheless HdCD19+ cells show slightly less proliferative potential compared to ACPA+RACD19+ when co-culture, as shown in figure 5.6 but this can be due to the small sample size.

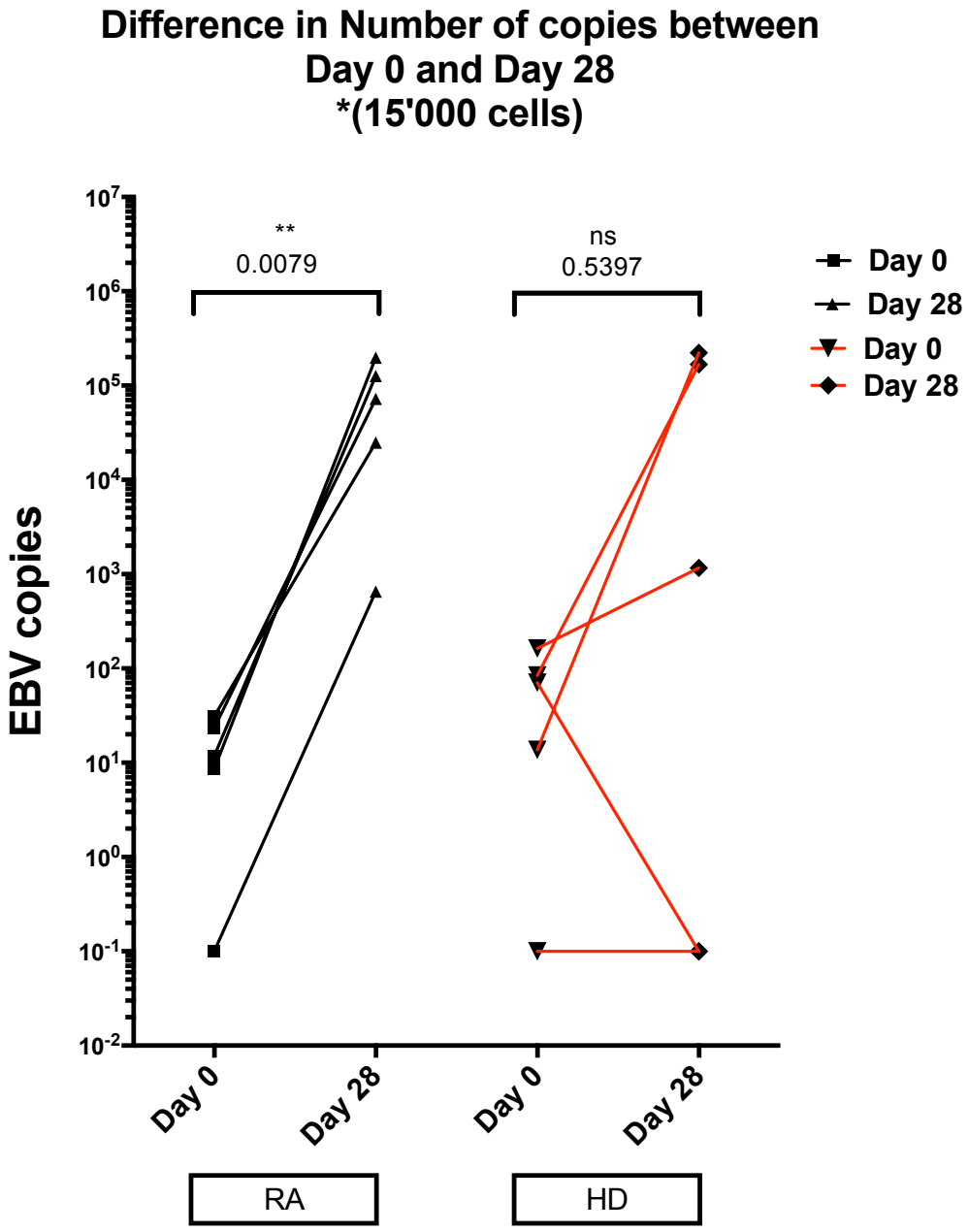


Figure 5.6. Co-culture of RA-FLS with RACD19+ and HdCD19+ B cells shows increase of EBV+ infected cells: Comparison of number of EBV copies detected between RACD19+ (5 peripheral blood samples) and HdCD19+ (5 peripheral blood samples) co-cultured with RA-FLS at day 0 and day 28. After 28 days

EBV-driven proliferation of B cells happens almost exclusively when cultured with RA-FLS. Nevertheless HdCD19+ cells show slightly less proliferative potential compared to ACPA+RACD19+. Samples that have not given any qPCR signal have been plotted as 0.

Increase in immunoglobulin and inflammatory cytokines production in presence of EBV -driven proliferation

ELISA tests on SN showed a strong increase of IL-6 and IL-8 in all proliferative co-culture, when compared to controls, both at 14 and 28 days (Fig 5.7). As for previous results, production of IL-6 and IL-8 does not exclusively increase in co-cultures of RA-FLS and RACD19+ but also in all HdCD19+ experiments in which proliferation has been detected (fig 5.7). There is no significant difference between the RA and HD group neither at 14 nor 28 days.

Production of IgM and IgG follows a similar pattern, with IgG production slightly more marked in RA (Fig. 5.8). Nevertheless, these differences – both at 14 and 28 days – have resulted not significant between the RA and HD group.

At least in one sample, considering the highest point reached in terms of production, I estimated an average of 34.50-fold increase for IL-6 and 320.76-fold for IL-8, compared to controls.

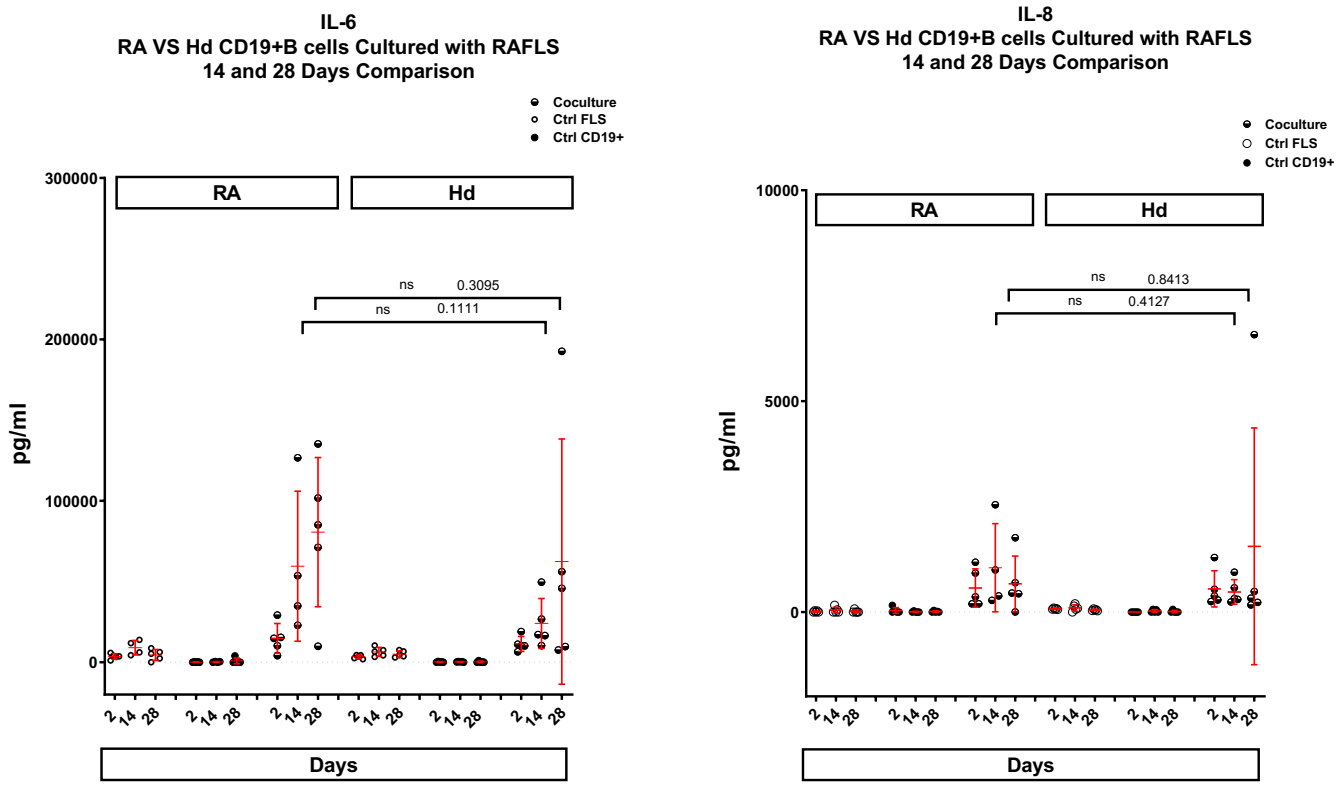


Figure 5.7. Increase of IL-6 and IL-8 production by co-culture system in the presence of EBV-driven proliferation: Results are showing Il-6 and Il-8 increased production when EBV driven proliferation is observed in co-culture of RACD19+/HdCD19+ with RA-FLS (timepoint: day2, day14, day 28). Increase is already appreciable at 14 days (Data from experiment comparing 5 RACD19+ peripheral blood samples and 5 HdCD19+ peripheral blood samples co-cultured with RA-FLS).

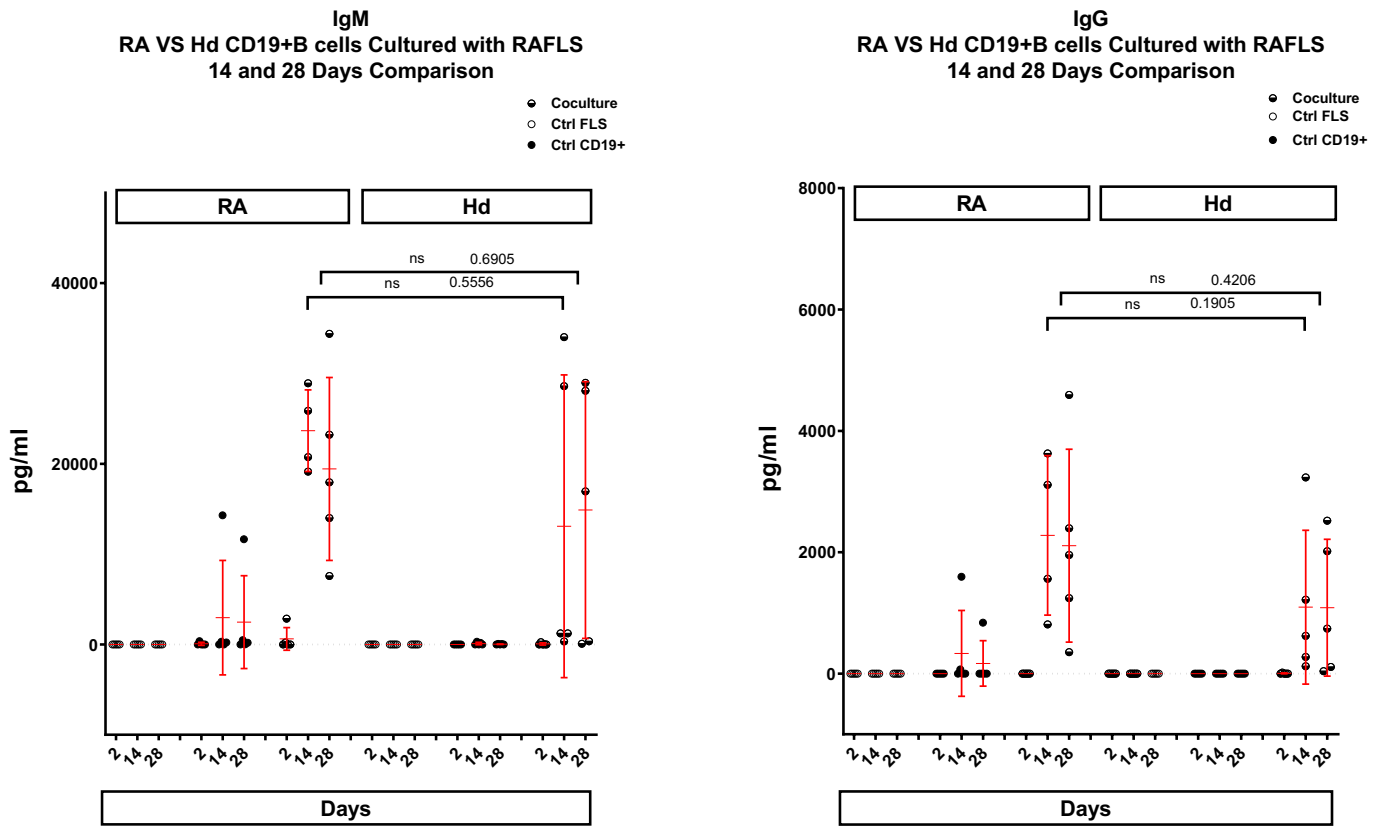


Figure 5.8. Increase of IgM and gG production by co-culture system in the presence of EBV-driven proliferation: Results are showing IgM and IgG increased production when EBV driven proliferation is observed in co-culture of RACD19+/HdCD19+ with RA-FLS (timepoint: day2, day14, day 28). Increase is already appreciable at 14 days (Data from experiment comparing 5 RACD19+ peripheral blood samples and 5 HdCD19+ peripheral blood samples co-cultured with RA-FLS).

Increase of CD38+ and IgM+ and switched IgG+ subsets during co-culture in presence of EBV-driven proliferation

I demonstrated by FACS analysis that in co-culture – during the EBV driven proliferation - there is a significant increase of proliferating CD38+ plasmablast. This increase of CD38+ is not present in ctrl B cells, so it is a further demonstration that this subset it is not affected by the stimuli. Also, IgM+ are significantly increased in co-culture, but a smaller - but still significant – increase is present also in the Ctrl B cells.

A smaller, but not significant, increase of switched IgG has been detected in co-culture compared to control (Fig 5.9).

Nevertheless, there is no significant difference about these markers when comparing co-culture with RACD19+ versus the HdCD19+ group (data not shown).

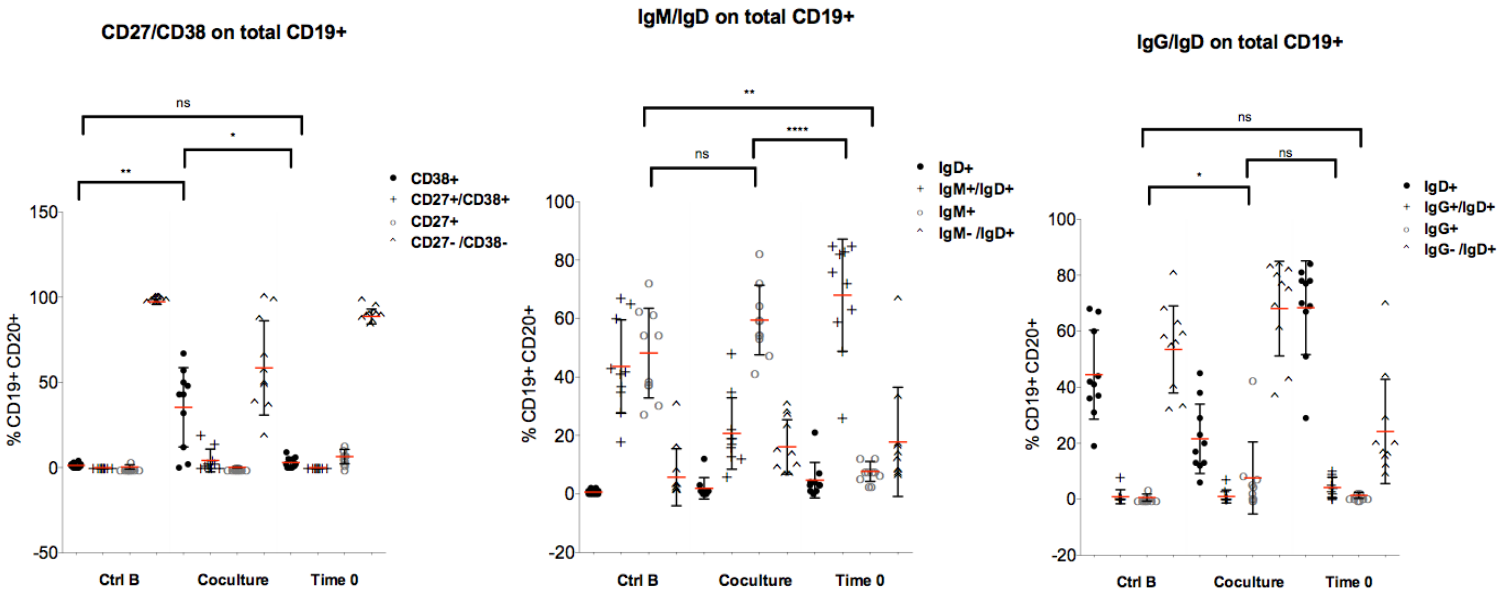


Figure 5.9 Increase of CD38+, IgM+ and switched IgG+ subsets during co-culture in presence of EBV-driven proliferation: Facs Results are showing CD38+ and IgM+ significant increase in proliferative co-culture. A milder increase of IgG has been detected. Graph are obtained from total CD19+CD20+ sorted B cell at the end of the co-culture. Result are showing percentages (Data from experiment comparing 5 RACD19+ peripheral blood samples and 5 HdCD19+ peripheral blood samples co-cultured with RA-FLS) .

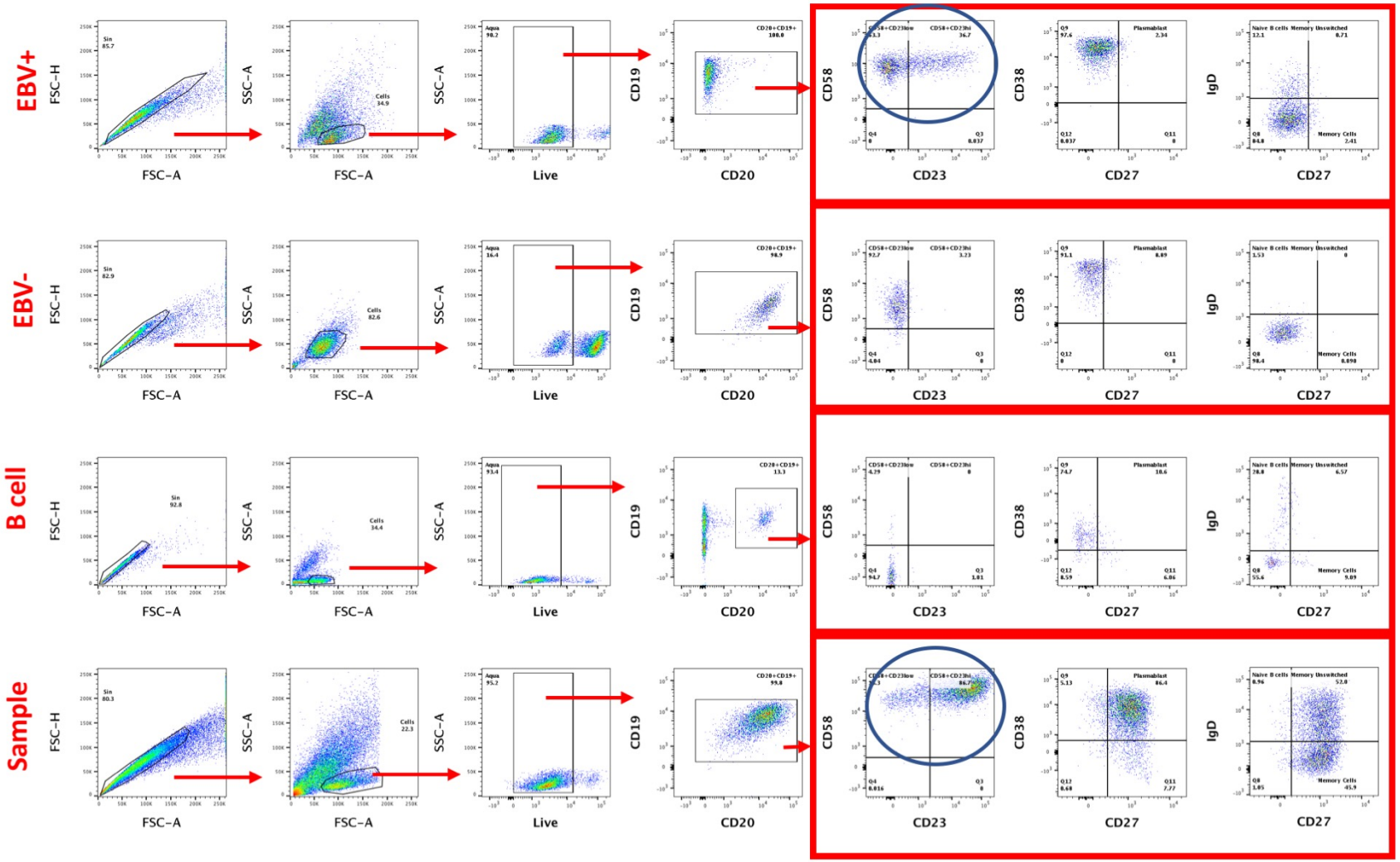
CHAPTER 6 Specific patterns emerging in RACD19+ and RA-FLS co-culture when EBV is detected

During experiment 0051 not all sorted B cells were needed to perform the qPCR, so the extra amount has been used to perform further experiment. These 30'000 "leftovers" lymphoblasts have been kept for 7 days in CM-B and then plated again a second time on RA-FLS without any kind of stimuli. Surprisingly cells were still proliferating at the end point of the second co-culture (day 36) (Fig. 3,1). Supernatant has then been collected every week and cells have been used to test reliability of the new CD58/CD23 Lymphoblast panel.

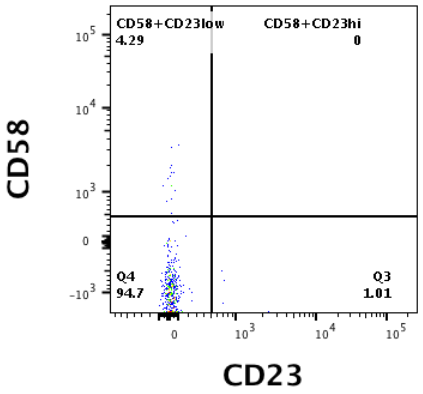
Further characterization of subsets of EBV+CD19+ B cells obtained from co-culture

Proliferating cells in co-culture are mainly CD58+/CD23^{low} and CD58+/CD23^{hi} B cells

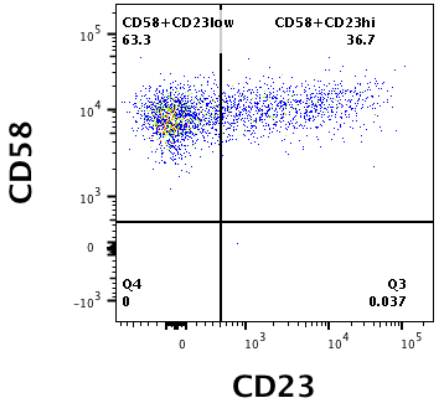
At the end of the experiment, all 0051 B cells have shown to belong to the main subsets described by Megyola, CD58^{+}/CD23^{low} and CD58^{+}/CD23^{hi} which are almost totally absent in a healthy control [205]. Also, compared to the positive control (Namalwa cells) my population appear highly skewed toward the proliferating compartment (CD23^{hi}) (Fig 6.1, Table 3.2). Interestingly, most of the co-culture at time 0 appear to have more CD58/CD23^{low} and very few CD58+/CD23^{hi} (Fig.6.2)


EBV+
EBV-
B cell
Sample
FSC-H
FSC-A
SSC-A
Live
CD19
CD20
CD58
CD23
CD38
CD27
IgD
Plasmablast
Memory Cells

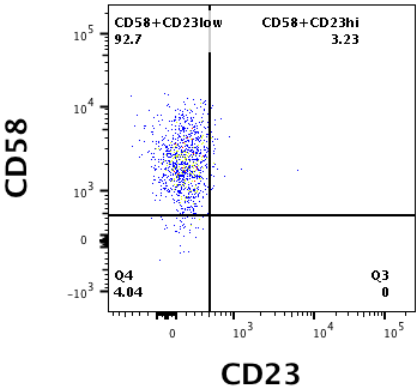
Hd PBMC



EBV+



EBV-



Sample 0051

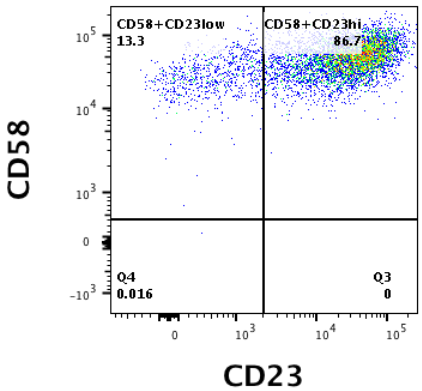


Figure 6.1. Expression of CD58 and CD23 marker by FACS in experiment 0051 (previous pages): Results are showing presence of the two previously described population $CD58^{+}/CD23^{low}$ and $CD58^{+}/CD23^{hi}$ in EBV infected cells from 0051, which are absent in healthy donor PBMC and EBV- cell line (Ramos cells). Nevertheless, the two subsets clearly resemble an EBV+ cell line (Namalwa). Gating has been done on singlets, lymphocytes, live cells and then expression of CD19 and CD20. Cell have been further characterised for expression of CD27/CD38 and IgD/CD27 to check presence of Naïve, Memory and Plasmablast cells.

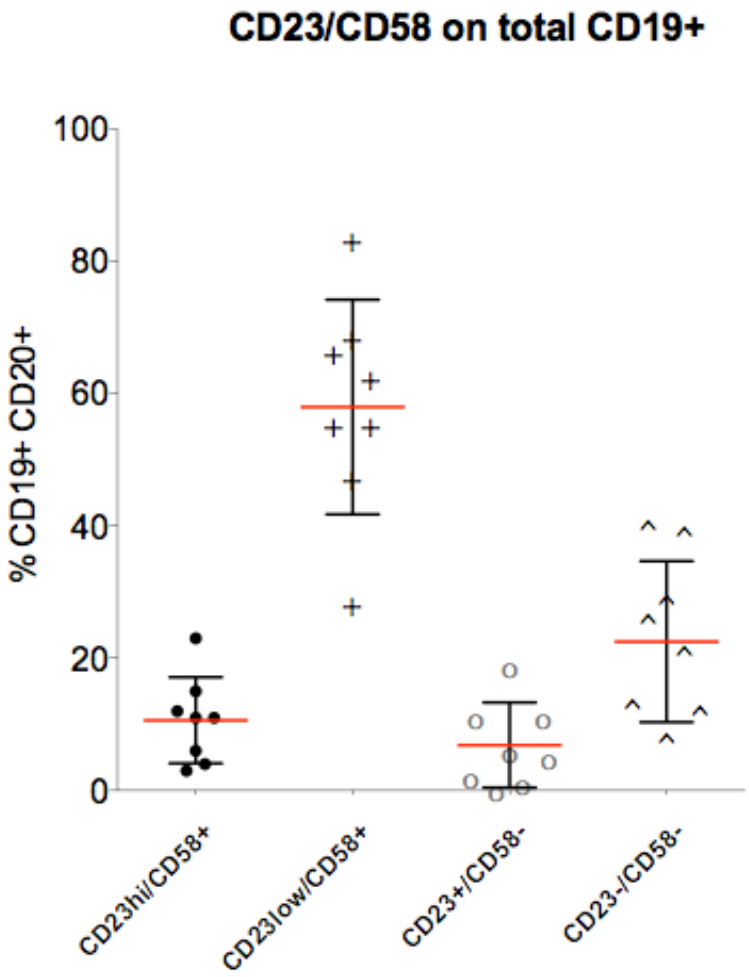


Figure. 6.2 Expression of CD58 and CD23 marker by FACS in all co-culture at time 0: Results are showing the expression for CD58 and CD23 markers at time 0 in co-culture, Interestingly, at the beginning of the experiment the majority of cells are $CD58^{+}/CD23^{low}$, while a very few $CD58^{+}/CD23^{hi}$ cells are present (Data from experiment comparing 5 RACD19+ peripheral blood samples and 5 HdCD19+ peripheral blood samples co-cultured with RA-FLS previously showed).

Screening trough multiplex reveal a different pattern in cytokines and chemokines produced between B cells and lymphoblast

Supernatants from experiment 0051 (first round: 28 day of co-culture; interphase with only LCL: 7 days ; second round of co-culture: 36 days) have been tested with a custom LEGENDplexTM assay panel and several different analytes were present (all analytes are present in table 6.1). In order to understand the mechanism and decide further approach analytes have been divided between those produced both in the presence of normal B cell and in the presence lymphoblast (Fig.6.3) and those produced only after lymphoblast proliferation (Fig 6.4). The entire panel analysed can be found in appendix (table 6.1)

The analytes produced in the presence of both cells are MIP-1alfa, CCL2, BAFF and IL-6.

The analytes produced only during lymphoblasts proliferation are CCL5, IL-12p70, TNFalfa and CXCL10.

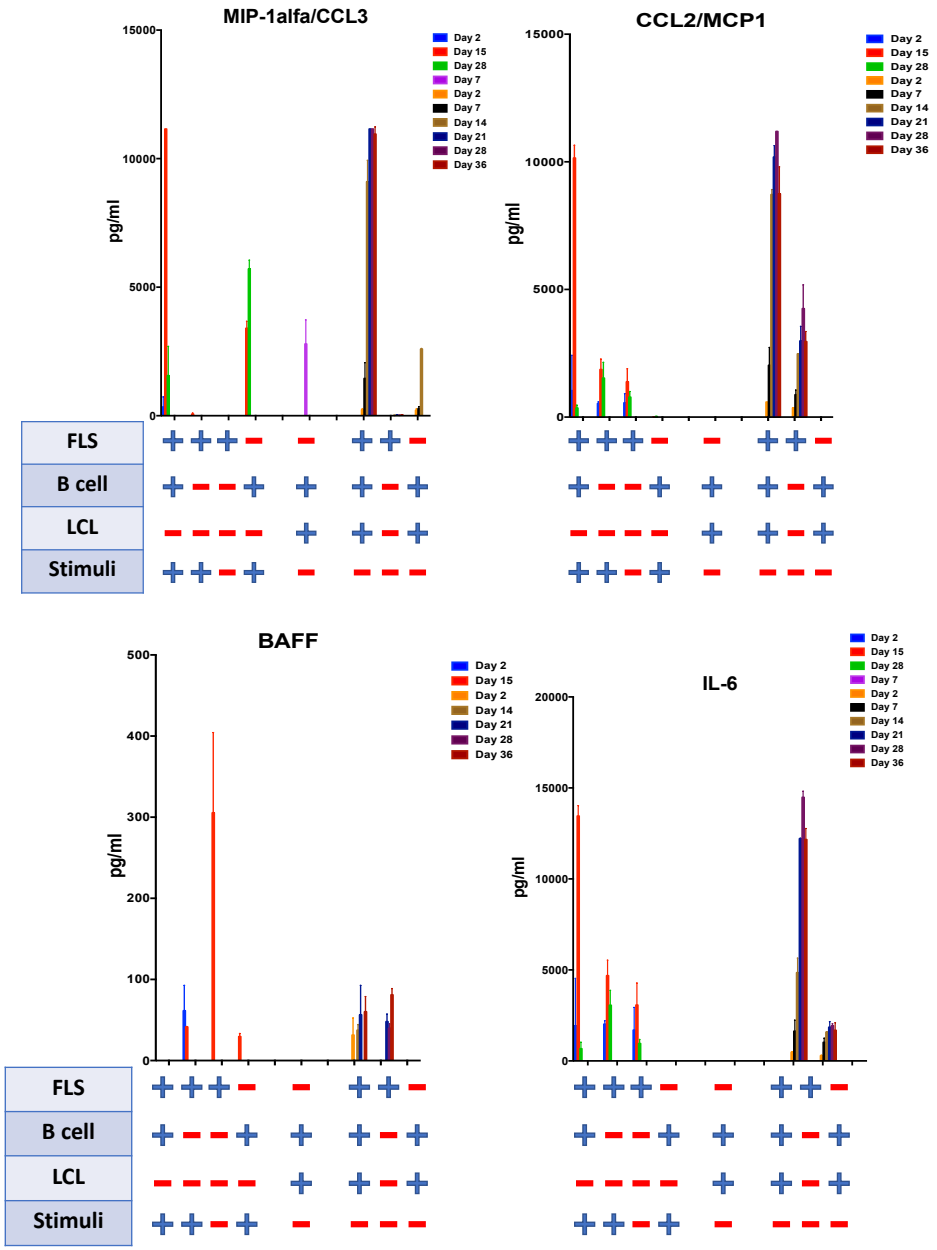


Figure 6.3. Analytes produced both in the presence of B cells and Lymphoblast: Results are showing analytes produced both in the presence of B cells and lymphoblast (data is showing results for only the single co-culture of RA-FLS 040/2016 and EMRRTB 0051).

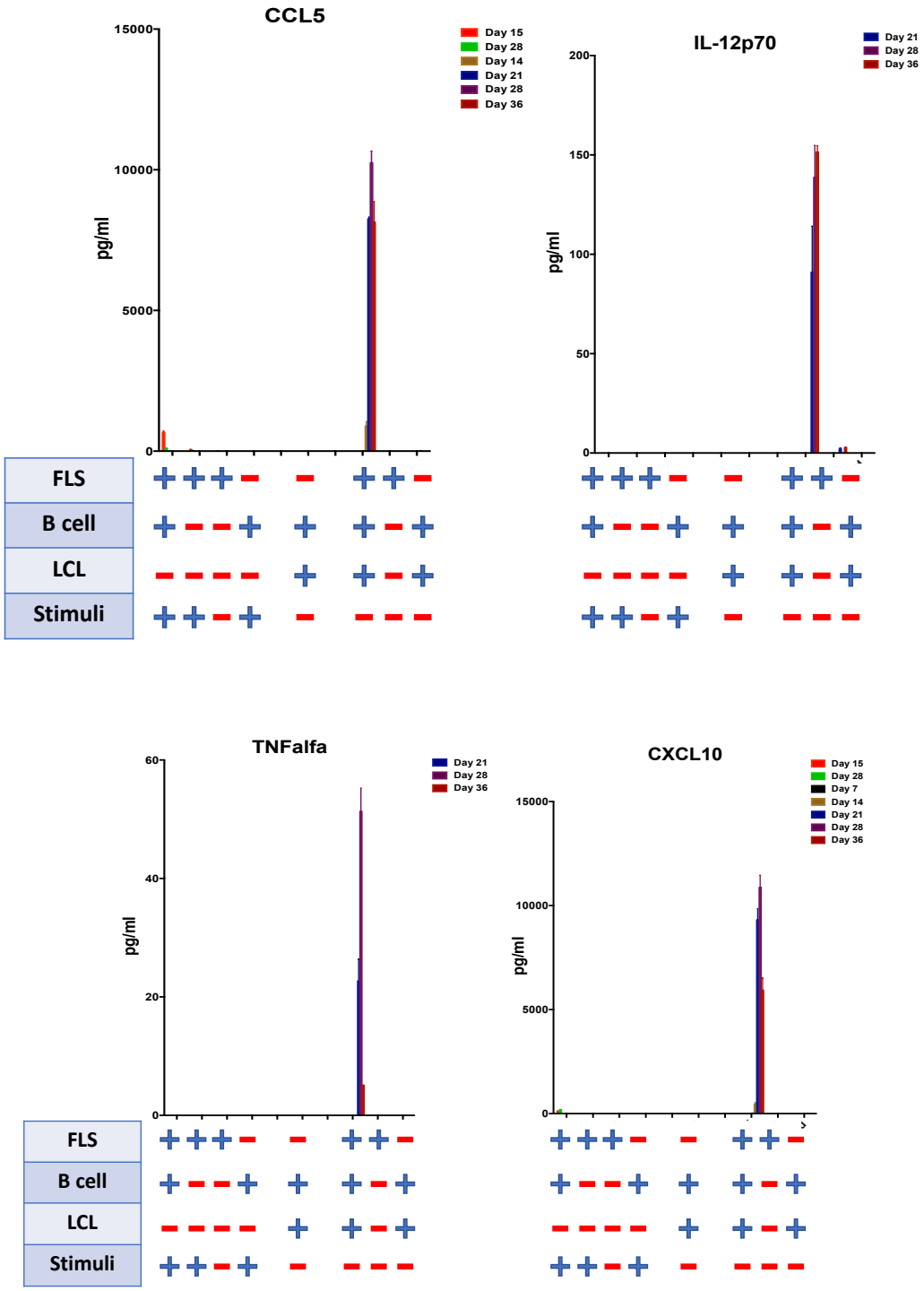


Figure 6.4. Analytes produced only in the presence of Lymphoblast: Results are showing analytes produced only in the presence of lymphoblast (data is showing results for only the single co-culture of RA-FLS 040/2016 and EMRRTB 0051).

CD58+/CD23^{hi} cells show an increased number of EBV EBNA-LP mRNA compared to CD58+/CD23^{lo} cells when analysed at single cell level

CD58+/CD23^{hi} and CD58+/CD23^{low} have been tested at single cells level for level of transcript from the BamHI W, repeated region: EBNA-LP. Intriguingly, CD58+/CD23^{hi} showed a significant increase in transcript signal when compared to CD58+/CD23^{low} (Fig.6.5). Since this compartment is the one proliferating the most it might be that is also the one with a stronger transformation potential.

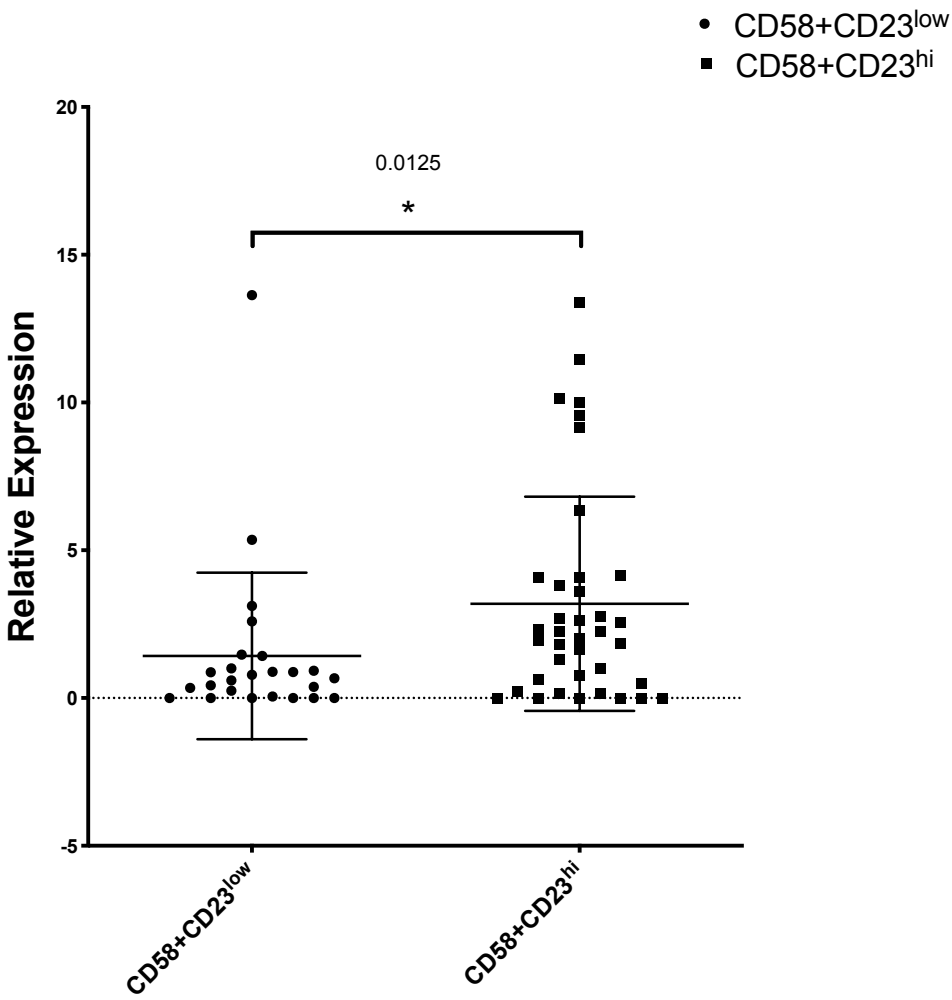


Figure 6.5. CD58+/CD23^{hi} cells show an increased number of EBV EBNA-LP mRNA compared to CD58+/CD23^{low} cells when analyzed at single cell level: Results are showing normalized relative expression values from qPCR for EBNA-LP RNA transcripts obtained from CD58+/23^{hi} and CD58+/CD23^{low}. values from CD58+/23^{hi} cells appear to be significantly higher than CD58+/CD23^{low}, meaning a higher content of the transcript. Number of cells with readable signal: CD58+/CD23^{low} = 25; CD58+/CD23^{low}= 38.

After sorting, subsets CD58+/CD23^{hi} and CD58+/CD23^{low} tend to merge again in a mixed population after 2 weeks in culture

CD58+/CD23^{hi} and CD58+/CD23^{low} have been sorted separately and the same number of cells has been cultured for two weeks, in order to understand stability of these two subsets. Nevertheless, at the end point, cells showed again the same mixed population of 23 ^{low} and 23 ^{hi} (data not shown). The only main difference has been in the number of cells obtained from the double positive cells, but this is not surprising since it is known to be the one with the most proliferative power. Also, cells have been sorted again, serial diluted and screened for copy number with the BamHI gDNA qPCR. The result of this last experiment has not showed any significant difference in the sorted groups (Fig.6.6).

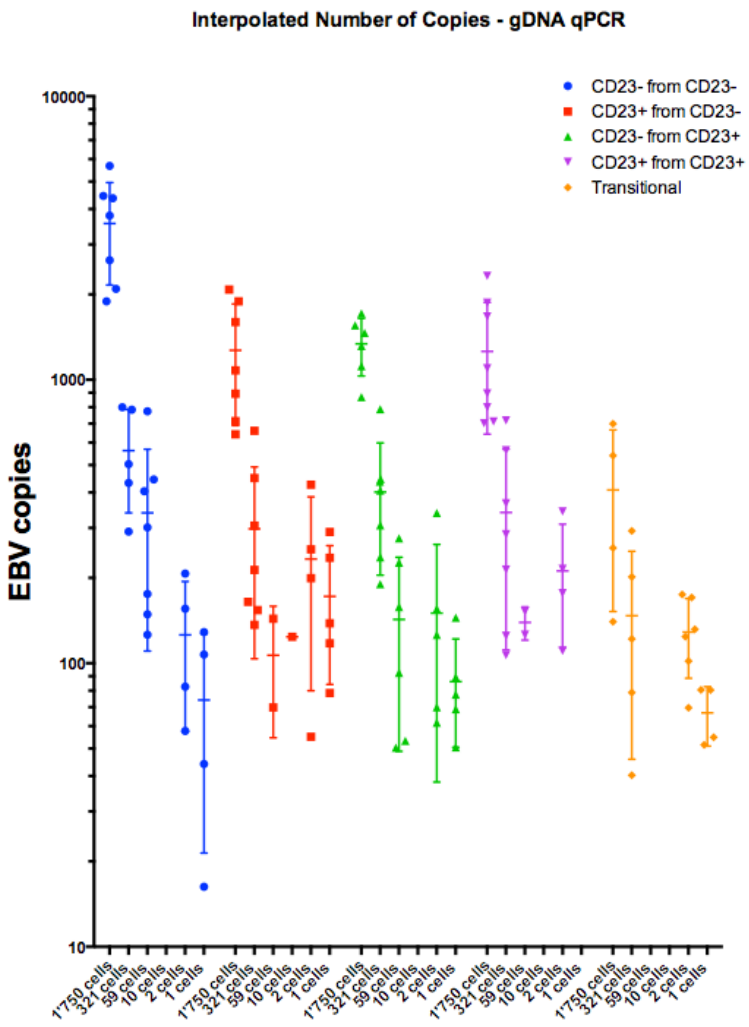


Figure 6.6. EBV Copies from separation of subsets CD58+/CD23^{hi} and CD58+/CD23^{low}: Results are showing BamHI gDNA qPCR interpolated copies number from sorted population of CD58+/CD23^{hi} and CD58+/CD23^{low}. Merged subsets have been subsequently separated in relative CD23^{low} and CD23^{hi} in order to perform the qPCR. No significant difference has been identified. *transitional cells are intended those on the limit between the two subsets.

A putative lymphoblastoid cell line isolation: Carejavi

Singles cells 10X genomics chromium sequencing and STRING analysis on lymphoblasts 0051 show presence of one main clonotype and typical proliferating blast phenotype

Lymphoblasts from experiment 0051 were expanded without stimuli or a feeding layer of RA-FLS for 4 months. In order to get a better understanding of these cells, I decided to perform an analyse using 10X genomic chromium single cell sequencing platform for both V(D)J and 5' total transcriptomic analysis. Ig sequences of the heavy chain are determined by a specific V(D)J rearrangement which defines a unique CDR3. Further diversity is introduced by N nucleotide addition during gene rearrangement and by point mutations through the process of SHM during clonal expansion of the B cells. The latter event brings to the formation of clonally related B cells. Thus, the Ig gene of these clonally related B cells derives from a single germline rearrangement and is characterised by the same V, D and J gene usage. Analysis of a set of IGH sequences from the lymphoblasts (experiment 0051) showed that almost all the cells were clonally related with a prevalence of one clonotype (Fig 6.7) This clonotype (isotype: IgM) was identified by Ig VH3-7, D2-8, JH4, CDR3=CARHRDQWAHW and Ig VL2-14, JL2, CDR3 = CSSYISNNILPVVL. Stability of this clonotype over time was confirmed after three months using a single cells nested PCR approach set up in my lab by Corsiero (data not shown) as alternative method to the 10X platform [110] .
Moreover, 5' total transcriptomic analysis showed that the three main clusters obtained by t-distributed stochastic neighbour embedding (t-SNE) shared most of the same upregulated genes (Fig.6.8). With STRING analysis was possible to confirm that the first 50 upregulated genes were related to process of cells undergoing active proliferation, i.e. DNA proliferation, chromosome and chromatin organization, DNA packaging, nucleosome organization. These are typical process present in proliferating lymphoblast (Fig 6.9). Of great interest is the fact that these cells have a strong signature for HLA-DRB1, the gene

connected to the shared epitope, and this might finally help me to connect these cells to an autoimmune mechanism (Fig.6.10).

As a candidate for a possible cell line, these cells have been provisionally referred to as “Carejavi”. (characterization and potential commercial exploitation of this cell line is included in future plans). Since these cells are from an ACPA+ RA patient, naturally harbour the virus and have a strong signature in HLA-DRB1 they might be a powerful tool to be used in research to understand the role of EBV in RA.

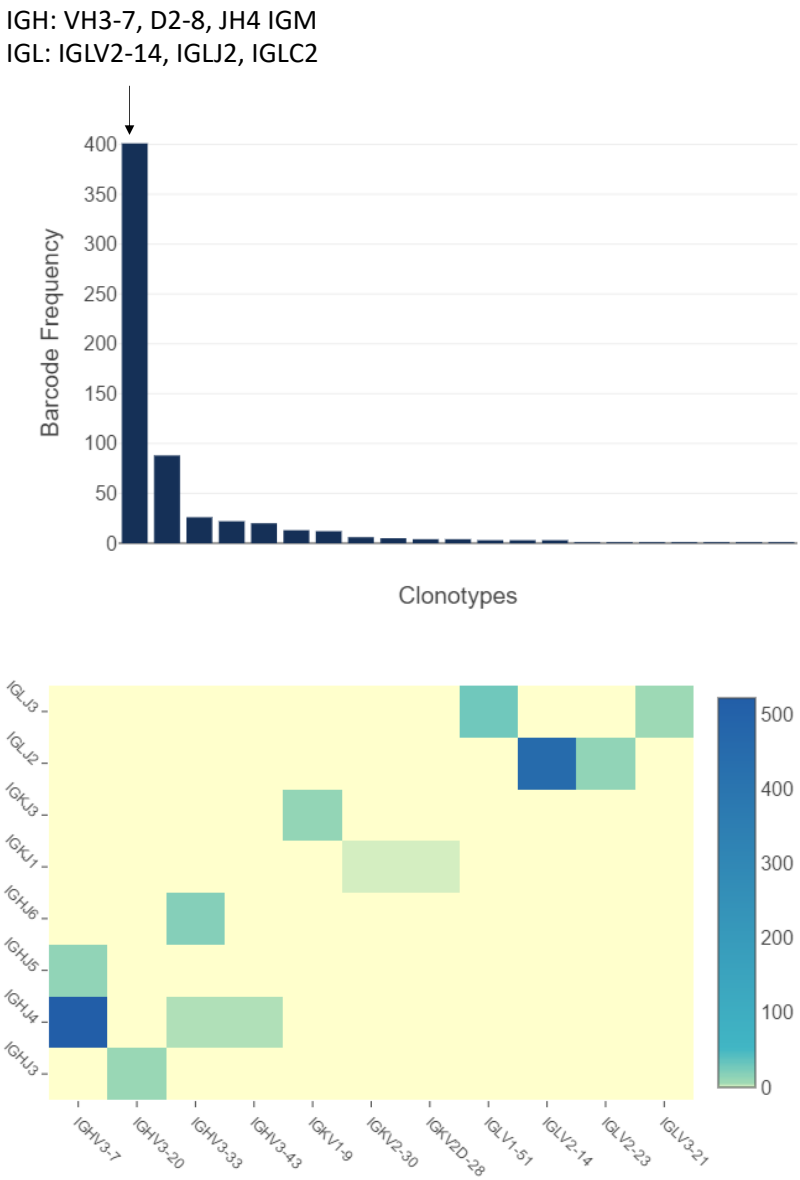


Figure 6.7. Barcode frequency (top) and heatmap (bottom) from 10X genomic single cells sequencing for V(D)J: Results are showing that almost all the cells analysed belong to a single clonotype (IgH: VH3-7, D2-8, JH4; IgL: IGLV2-14, IGLJ2).

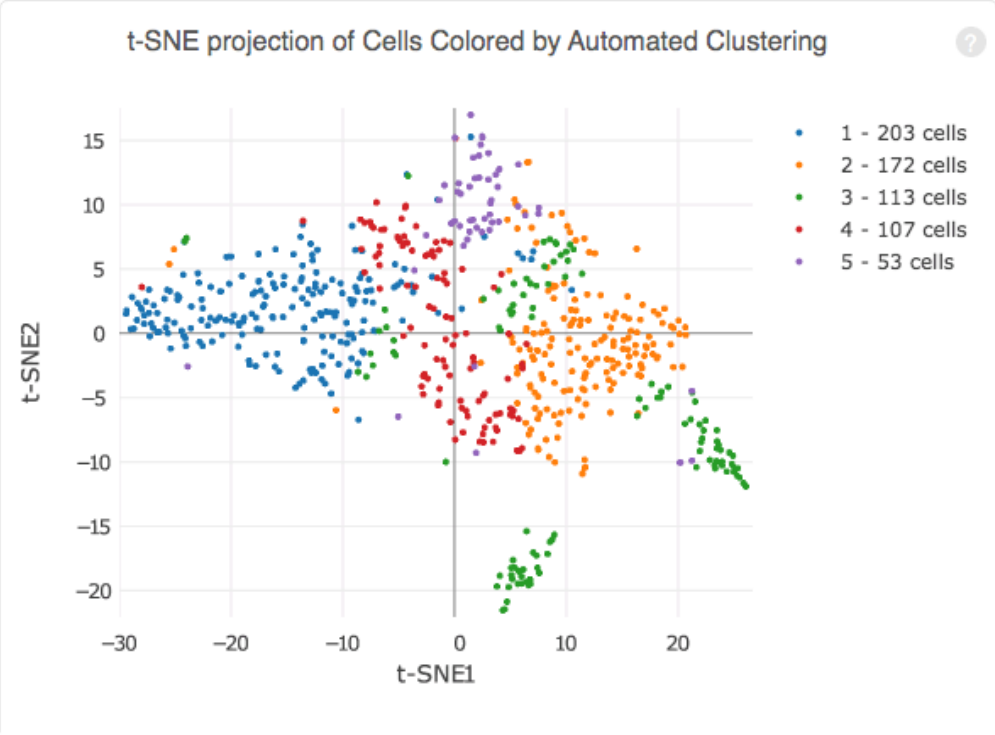
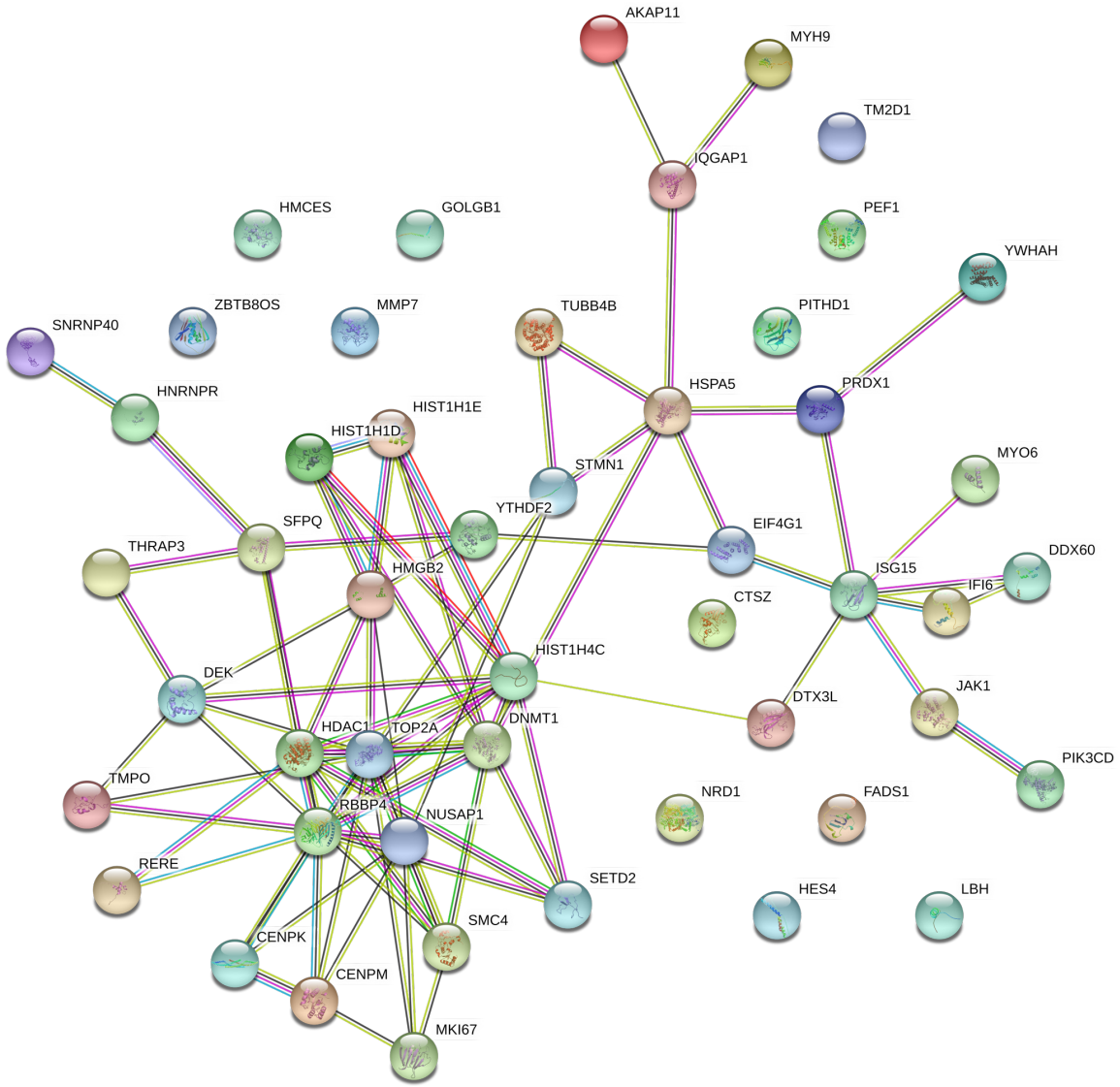
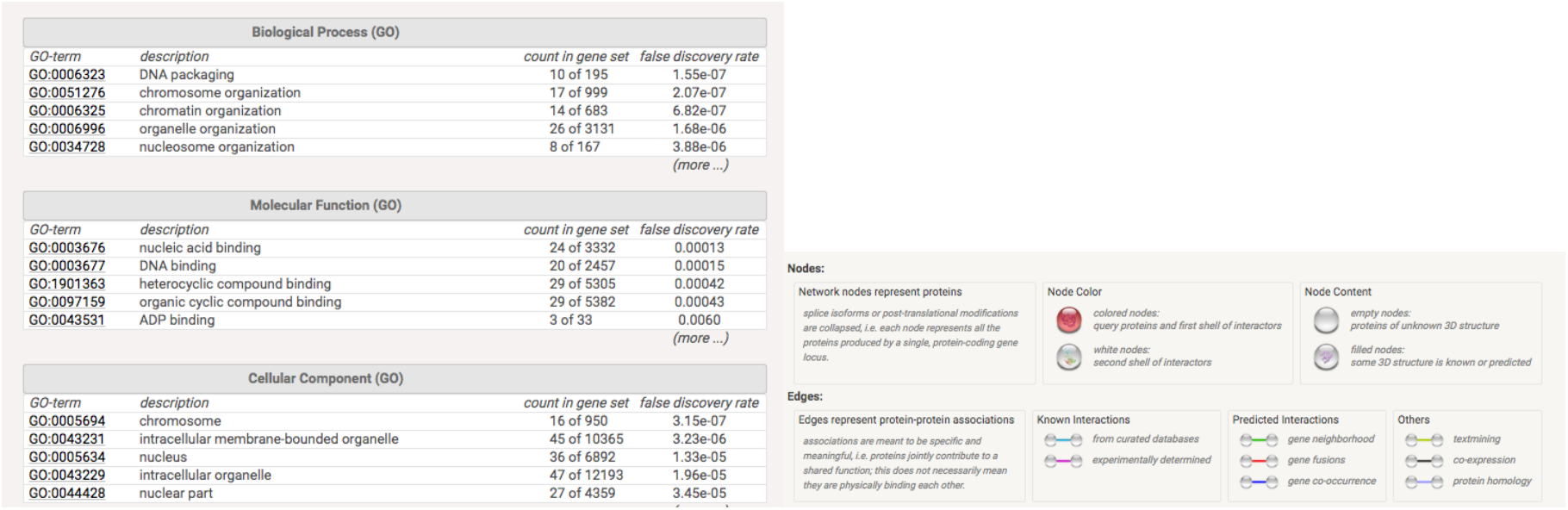


Figure 6.8. t-SNE from 10x genomic single cells sequencing for 5' transcriptomic: Results are showing 3 main cluster (1,2 and 3) obtained from 10 X t-SNE analysis. Cluster 4 and 5 have been excluded from the analysis due to the presence of high background genes, such as ribosomal or mitochondrial genes). From 10X genomic: *These are the assignments of each cell-barcode to clusters by an automated clustering algorithm. The clustering groups together cells that have similar expression profiles. The axes correspond to the 2-dimensional embedding produced by the t-SNE algorithm. In this space, pairs of cells that are close to each other have more similar gene expression profiles than cells that are distant from each other".*


AKAP11
MYH9
TM2D1
IQGAP1
HMCES
GOLGB1
PEF1
YWHAH
ZBTB8OS
MMP7
TUBB4B
PITHD1
SNRNP40
HNRNPR
HSPA5
PRDX1
HIST1H1E
HIST1H1D
MYO6
STMN1
YTHDF2
SFPQ
EIF4G1
THRAP3
DDX60
HMGB2
ISG15
IFI6
CTSZ
HIST1H4C
DEK
JAK1
DTX3L
DNMT1
HDAC1
TOP2A
PIK3CD
TMPO
NRD1
FADS1
RBBP4
NUSAP1
RERE
SETD2
HES4
LBH
SMC4
CENPK
CENPM
MKI67



Biological Process (GO)

GO-term	description	count in gene set	false discovery rate
GO:0006323	DNA packaging	10 of 195	1.55e-07
GO:0051276	chromosome organization	17 of 999	2.07e-07
GO:0006325	chromatin organization	14 of 683	6.82e-07
GO:0006996	organelle organization	26 of 3131	1.68e-06
GO:0034728	nucleosome organization	8 of 167	3.88e-06
			(more ...)

Molecular Function (GO)

GO-term	description	count in gene set	false discovery rate
GO:0003676	nucleic acid binding	24 of 3332	0.00013
GO:0003677	DNA binding	20 of 2457	0.00015
GO:1901363	heterocyclic compound binding	29 of 5305	0.00042
GO:0097159	organic cyclic compound binding	29 of 5382	0.00043
GO:0043531	ADP binding	3 of 33	0.0060
			(more ...)

Cellular Component (GO)

GO-term	description	count in gene set	false discovery rate
GO:0005694	chromosome	16 of 950	3.15e-07
GO:0043231	intracellular membrane-bounded organelle	45 of 10365	3.23e-06
GO:0005634	nucleus	36 of 6892	1.33e-05
GO:0043229	intracellular organelle	47 of 12193	1.96e-05
GO:0044428	nuclear part	27 of 4359	3.45e-05

Figure 6.9. STRING analysis for first 50 upregulated genes in Carejavi (with legend, starting from previous page): Results are showing the output obtained from STRING analysis for the first 50 upregulate genes from Carejavi. Output network is showing more significant interaction than expected, meaning that my proteins have more interaction among themselves than what would be expected for a random set of proteins of similar size, drawn from the genome. This enrichment indicates that the proteins are at least partially biologically connected as a group.

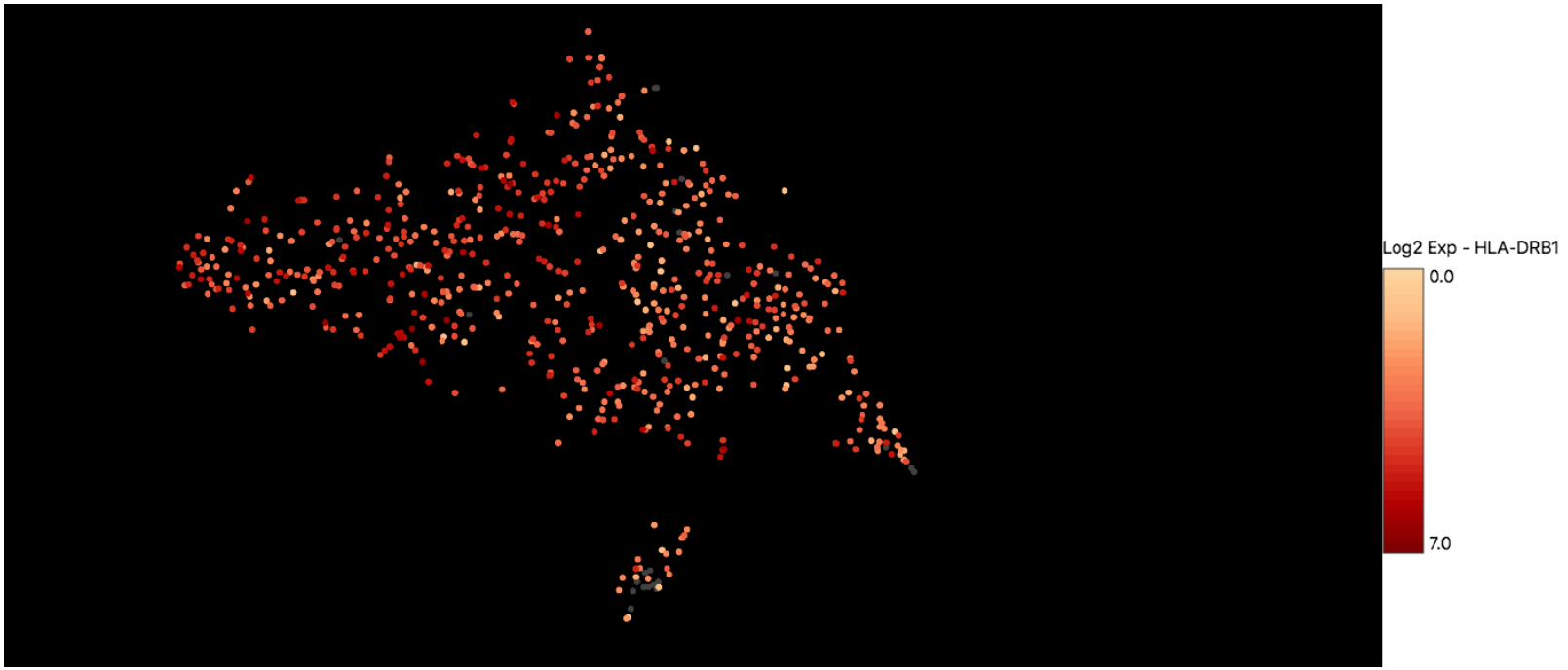


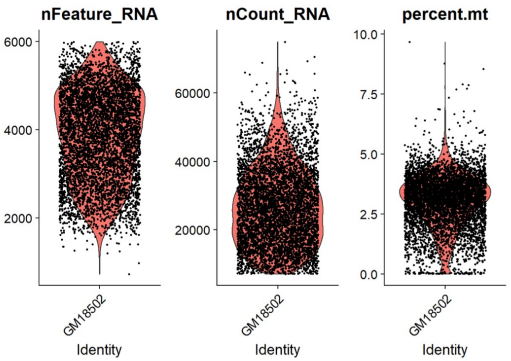
Figure 6.10. Expression of HLA-DRB1 in Carejavi Results are showing the HLA-DRB1 gene expression output obtained from 10X single cells genomic sequencing. Graphs has been obtained with 10X software LOUPE (10xGenomic)

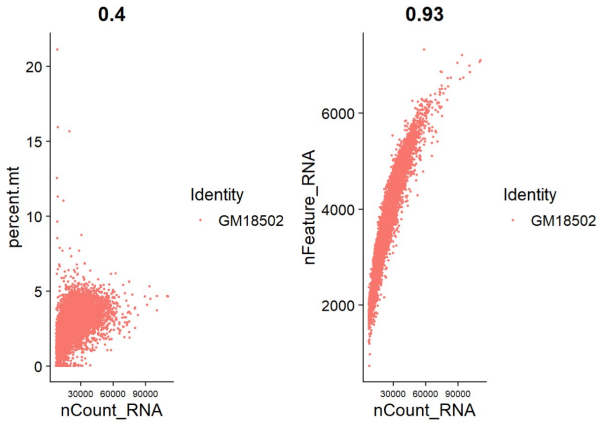
Seurat comparative analysis between Carejavi and two EBV+ cell line using online available single cells dataset: Carejavi express typical marker of LCLs.

An alternative analysis using Seurat has been performed on lymphoblasts from experiment 0051 in order to compare them with two publicly available single cell dataset of EBV+ cell lines: GM18502 and GM12878. These datasets have been obtained from Osorio et al [247]. The 3 cell lines have been analysed in parallel and showed similar QC results, in term of RNA feature (genes), RNA count (absolute count of RNA), percentage of mitochondrial genes (which account for apoptic cells) despite the difference in absolutes numbers (GM18502: 5437 cells; GM12878: 7186 cells; Carejavi: 943 cells (fig 6.11). Next, the identification of highly variable feature (feature that exhibit high cell to cell variation) has highlighted several similarities in the top 20 genes between the Carejavi and the other two cell lines, with CCL3, CCL4, and CCL22 as the most homogenously genes represented (fig 6.12). After regression of genes related to the cell cycle, this list of features has been then used to calculate PCs (including dimensional heatmap, elbow plot and jackstraw plot, Supplementary figure 6.1 - 6.4) in order to perform clustering and create a Uniform Manifold Approximation and Projection (UMAP) for non -linear dimensionality. For this analysis the first 10 PCs have been used (fig 6.13). The final expression heatmap of the 10 most representative genes has been generated for all the 3 cell lines according to the number of clusters and can be visualized in figure 6.14, 6.15 and 6.16. I particularly focused on the mutually exclusive compartmentalization of genes related to Immunoglobulin/differentiation versus activation/NFKBI (for this clustering list of genes suggested by Sorelle have been used [252], Fig.6.17-6.19). What is more interesting is that CD58 and CD23/FCER2 positive cells tends to be more closely clustered with the activation and NFKB targeted related genes. This is a further confirmation not just that the Carejavi are behaving like a

lymphoblastoid cell line but also that cells expressing CD23/FCER2 and CD58 are the one more prone to proliferation.

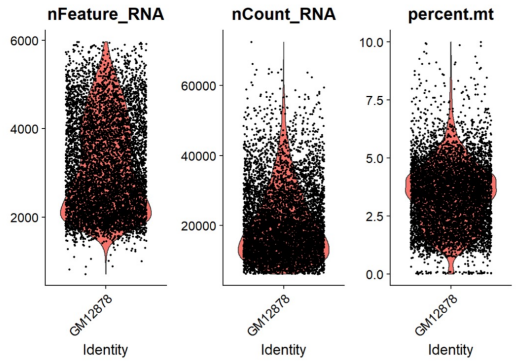
GM18502

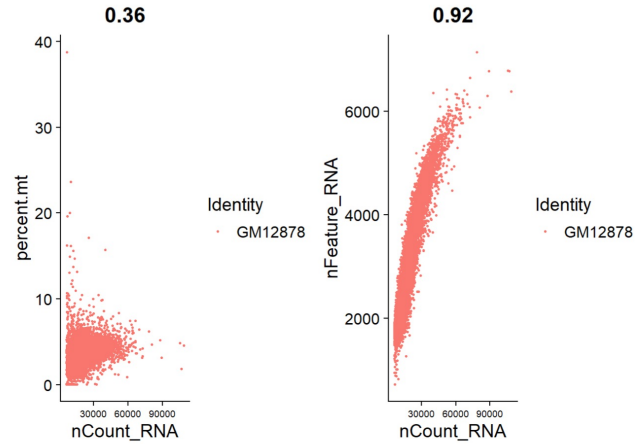




5437 cells

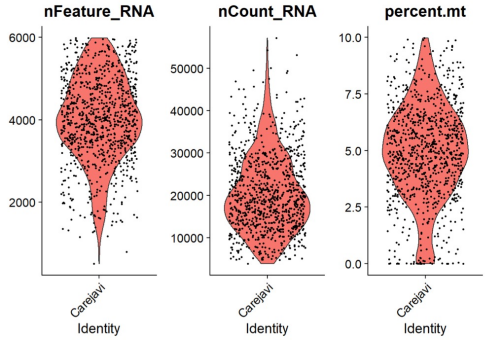
GM12878

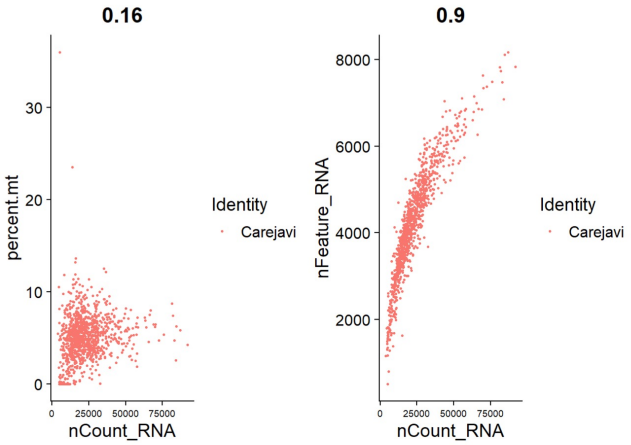




7186 cells

Carejavi





943 cells

(From previous page) Figure 6.11. Quality control results between the 3 cells lines: GM18502, GM12878 and Carejavi: Results are showing violin plots for RNA features (genes), count of absolute RNA molecules and mitochondrial genes percentage. Same data are expressed also as Scatter Feature plot. Absolute number of cells subset after QC is showed.

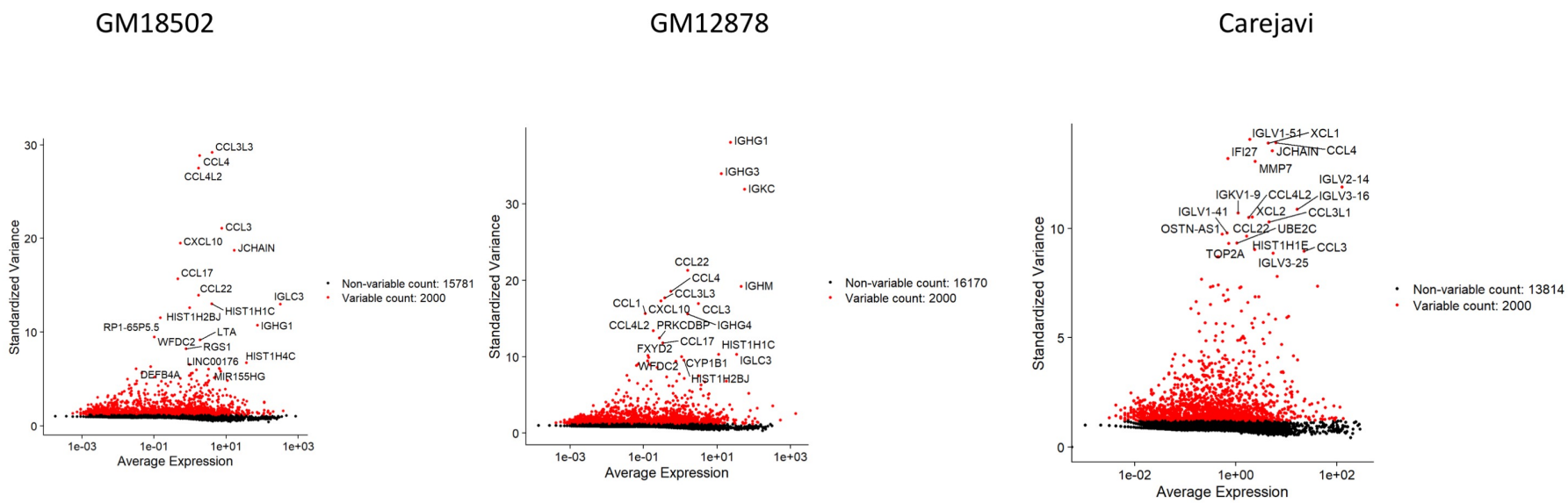


Figure 6.12. Identification of highly variable features between the 3 cells lines:GM18502, GM12878 and Carejavi: Results are showing features that exhibits a high cell to cell variation between the 3 samples.

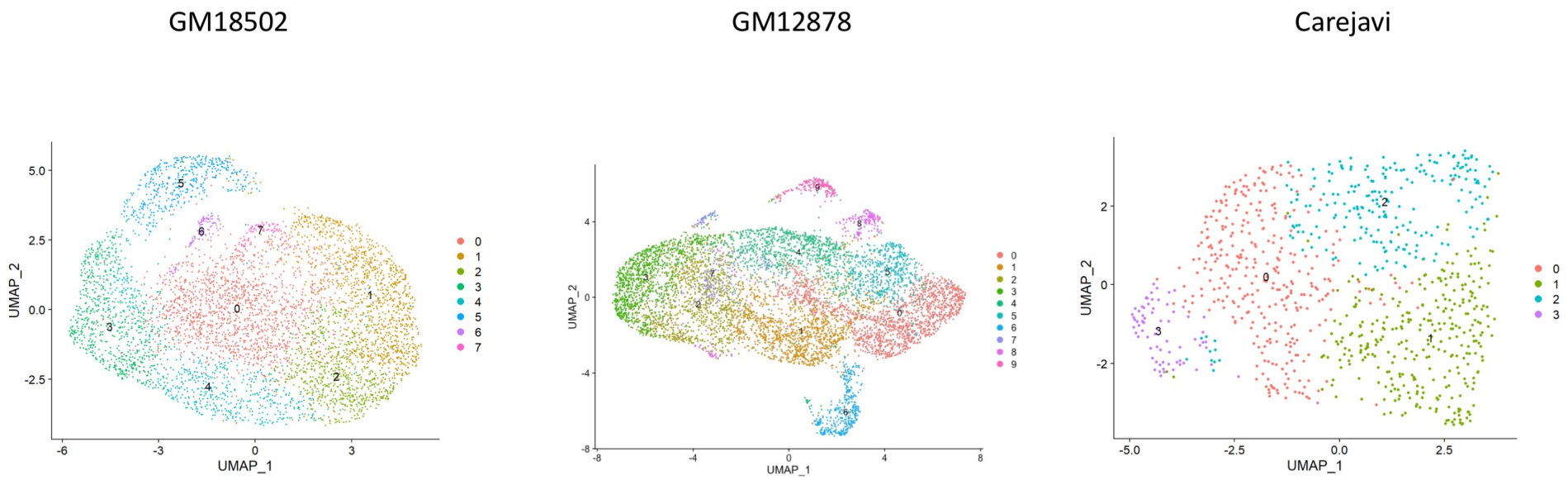


Figure 6.13. UMAP of the 3 cells lines:GM18502, GM12878 and Carejavi: Results are showing UMAP plot for 3 cell lines

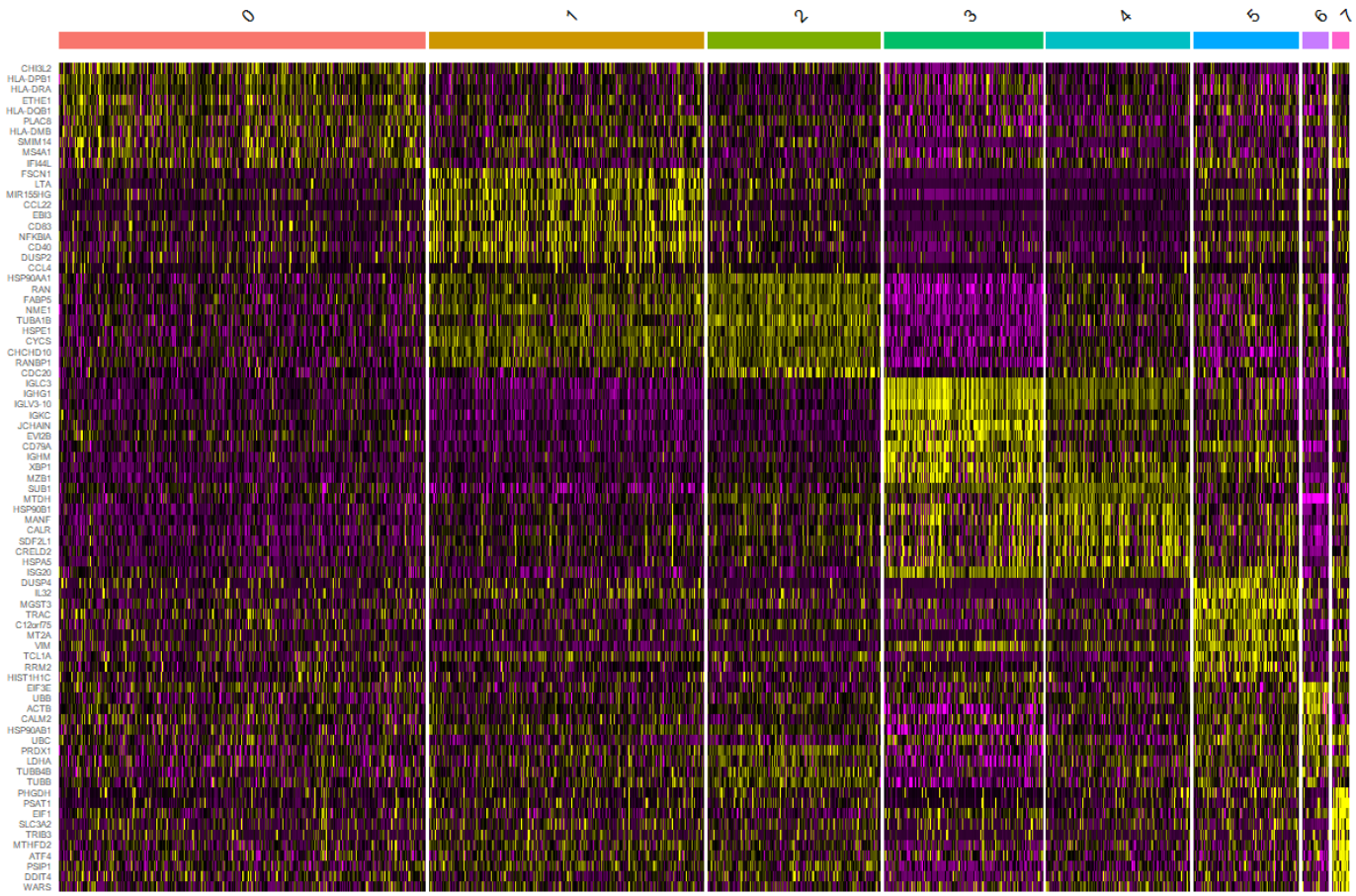


Figure 6.14. Expression heatmap for GM18502.

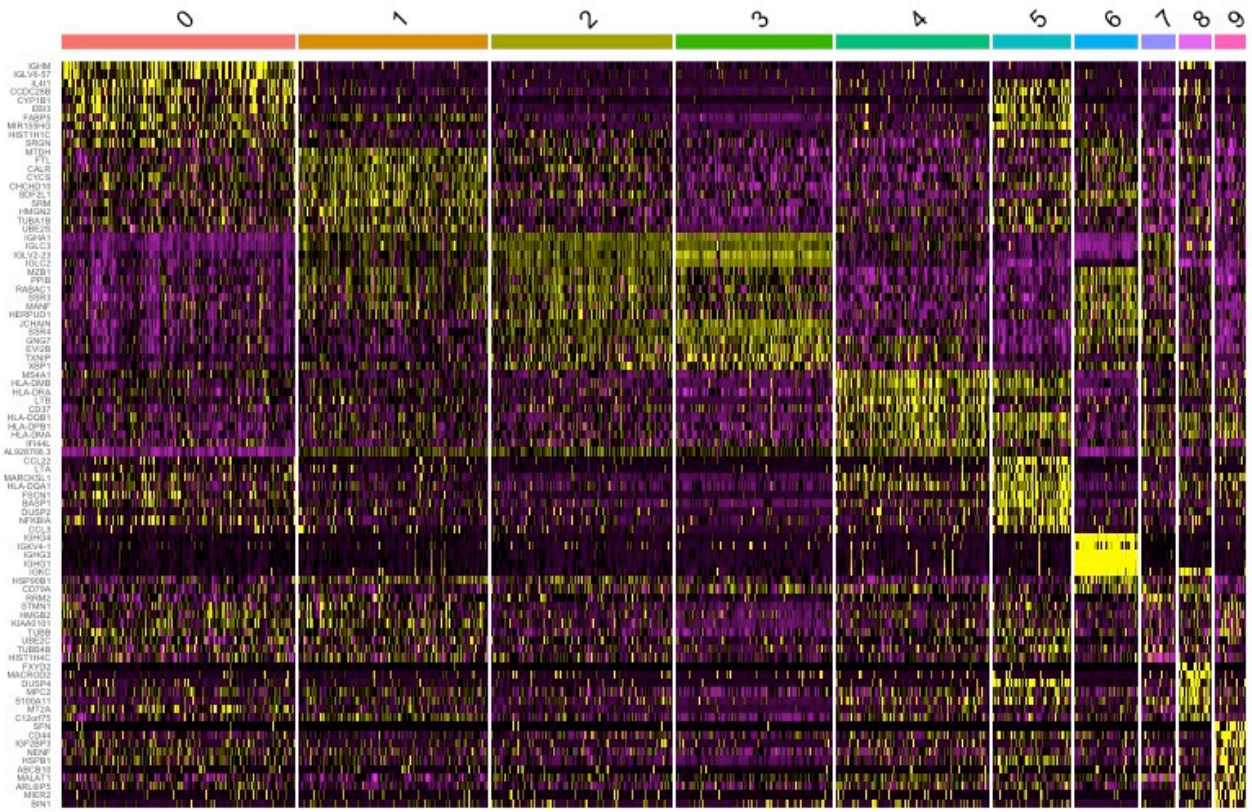


Figure 6.15. Expression heatmap for GM12878.

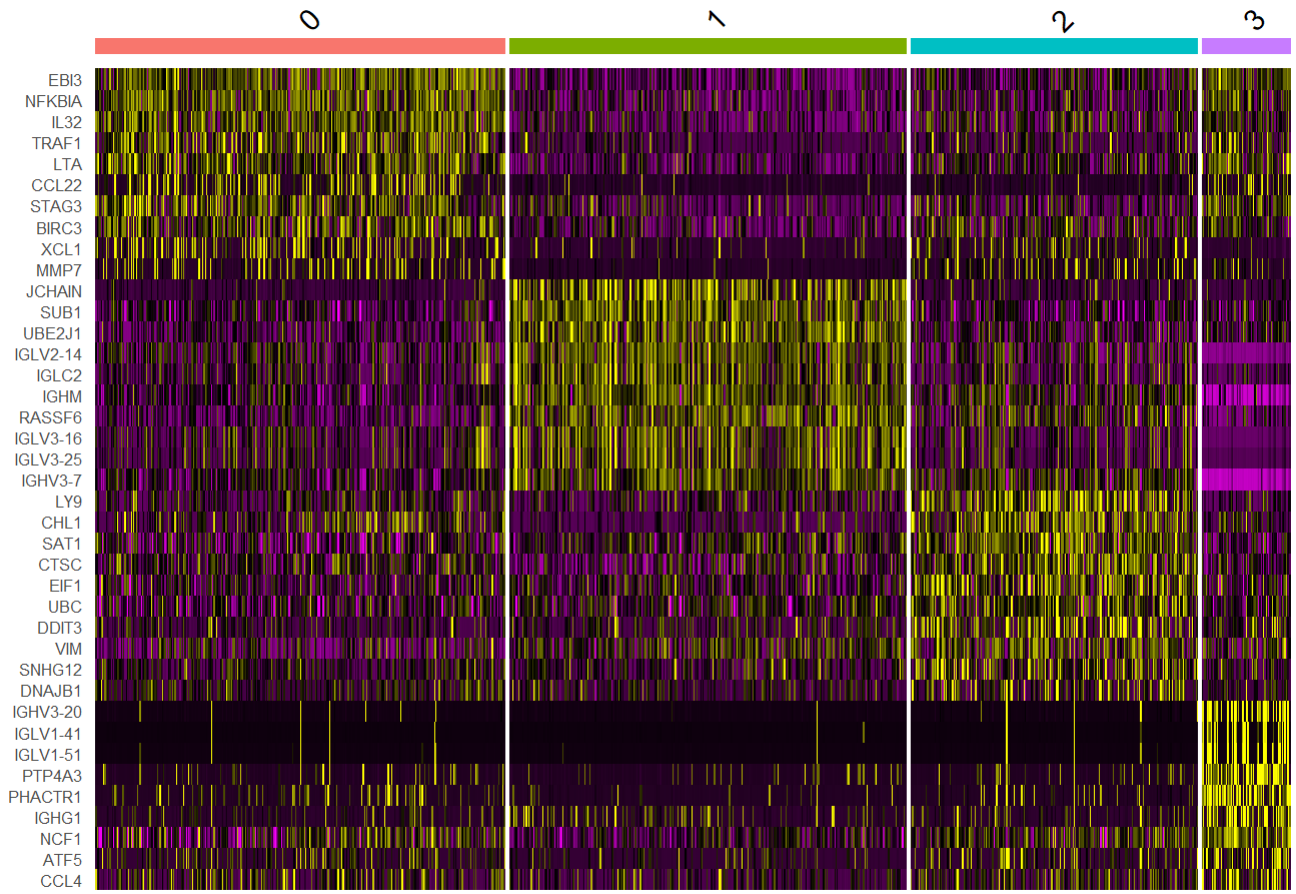


Figure 6.16. Expression heatmap for Carejavi.

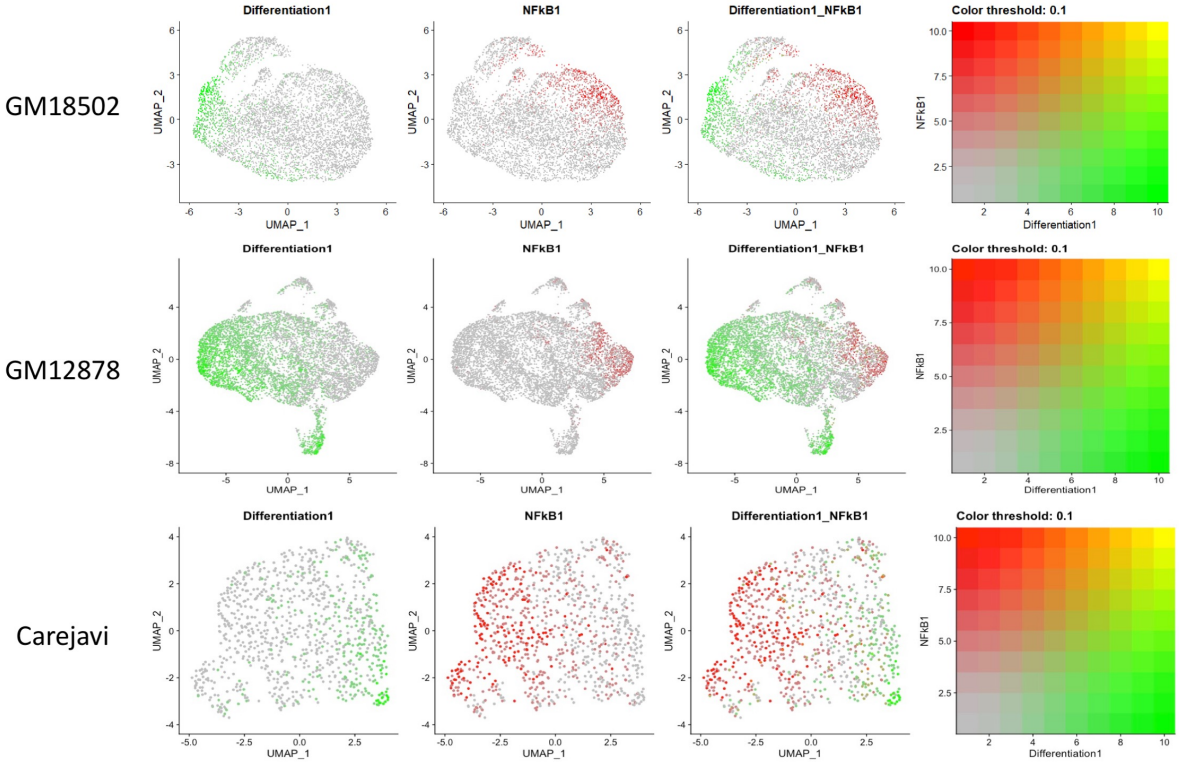


Figure 6.17. Comparison between the 3 cells lines for differentiation and NFKB related genes: Results are showing mutually exclusive clustering of cells expressing genes related to differentiation (TNFRSF17, XBP1, MZB1, CD27) and genes related to NFkB cascade and activation (NFKB2, NFKBIA, EBI3, ICAM1, BCL2A1) between the 3 different cell lines. Red is for NFkB signature, while green is for Differentiation signature.

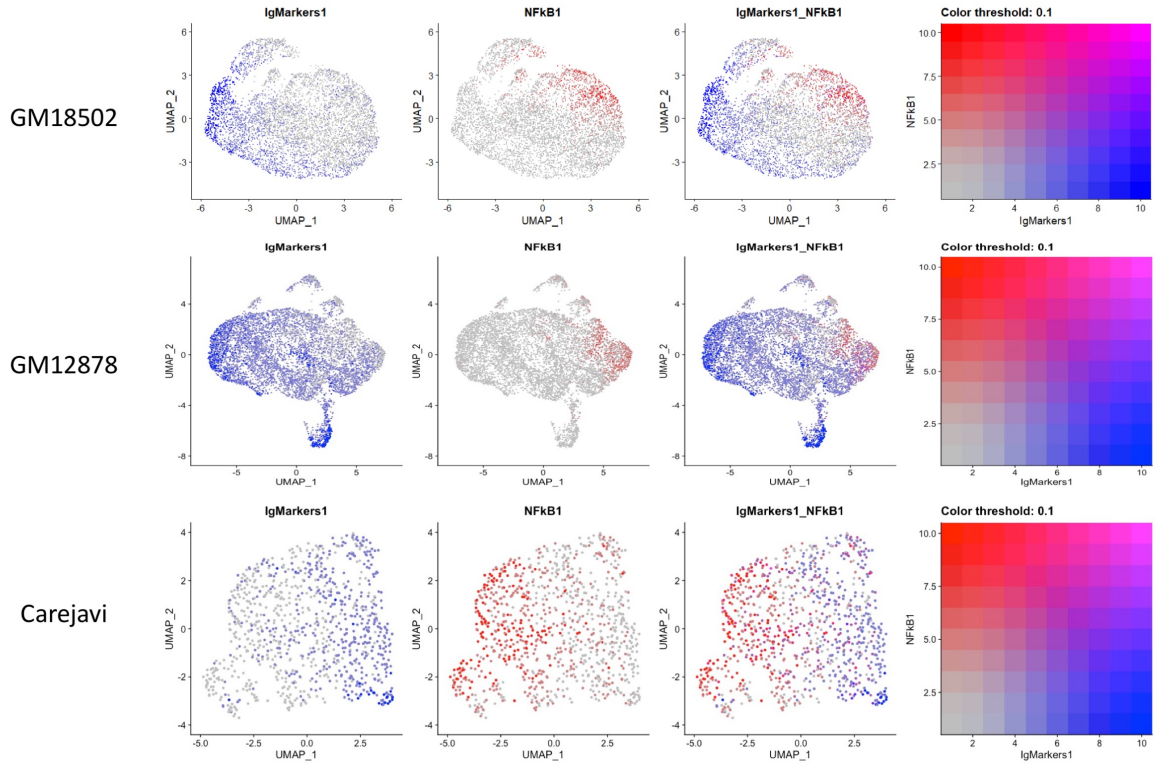


Figure 6.18. Comparison between the 3 cells lines for Immunoglobulin and NFKB related genes: Results are showing mutually exclusive clustering of cells expressing genes related to immunoglobulin (IGHM, IGHA1, IGHG1) and genes related to NFkB cascade and activation (NFKB2, NFKBIA, EBI3, ICAM1, BCL2A1) between the 3 different cell lines. Red is for NFkB signature, while blue is for Immunoglobulin signature.

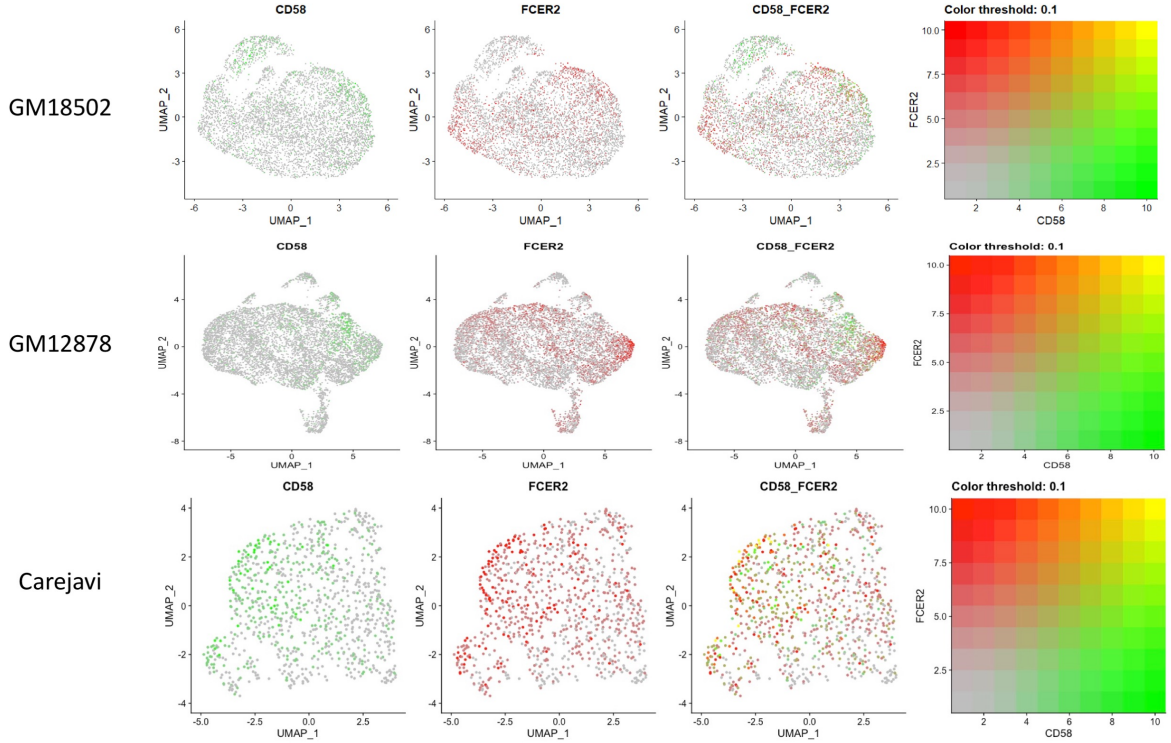


Figure 6.19. Comparison between the 3 cells lines for CD58 and CD23/FCER2: Results are showing cells expressing CD58 and CD23/FCER2 genes. Red is for CD23/FCER2 signature, while green is for CD58 signature.

Co-culture transcriptomic analysis: reciprocal effect of RACD19+ B cells on RA-FLS phenotype

Singles cells 10X genomic sequencing and STRING analysis reveal RA-FLS phenotype switch when in co-culture with RA-CD19+

One co-culture experiment (RA-FLS and RA-CD19+ B cells) and control (RA-FLS alone) were screened using single cells 10X genomic platform. Preliminary data showed that RA-FLS cultured with RA-CD19+ B cells tend to switch from a proliferative and invasive phenotype to a more inflammatory one **(**Fig.6.20, 6.21). Specifically, RA-FLS in co-culture showed to downregulate genes related to cell proliferation pathways, while upregulate those ones relative to type I interferon signalling, cellular response to organic substances and response to cytokine (Fig.6.22, 6.23).

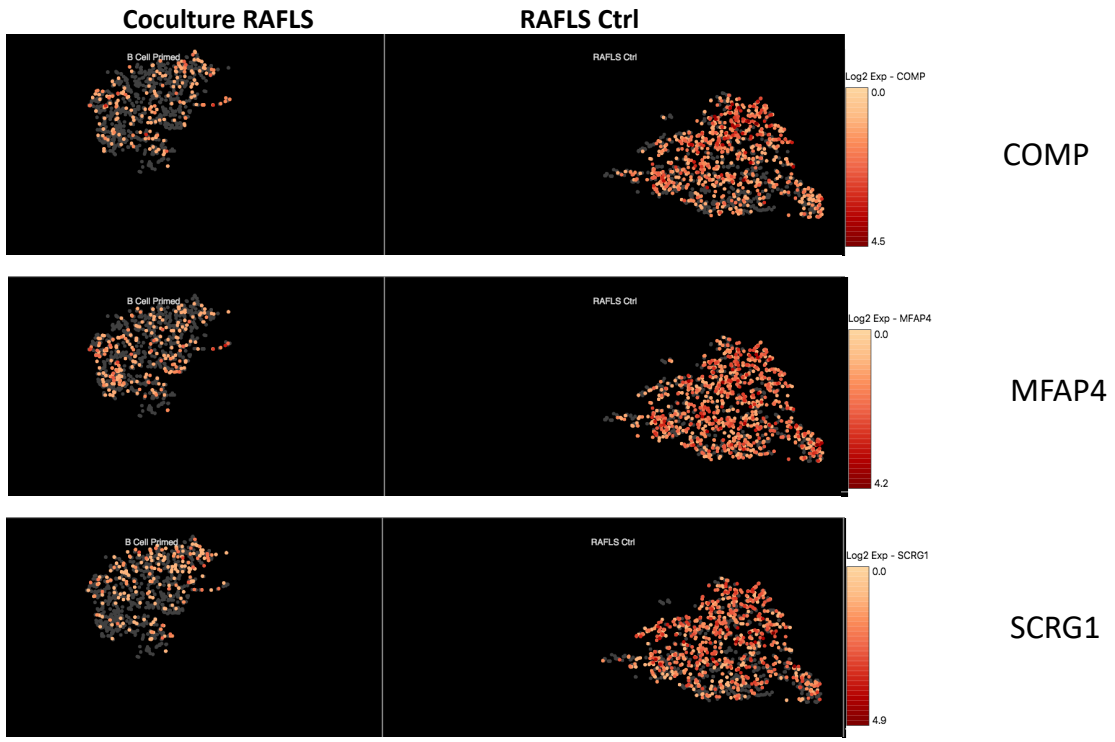


Figure 6.20. Differential gene expression between RA-FLS co-cultured with EBV+RACD19+ and control (downregulated) Results are showing that co-cultured RA-FLS (on the left) tend to downregulate expression of genes as Cartilage Oligomeric Matrix Protein (COMP), Microfibril Associated Protein 4 (MFAP4), and Stimulator Of Chondrogenesis (SCRG1). Relative expression output obtained from 10X single cells genomic sequencing. Graphs has been obtained with 10Xsoftware LOUPE (10x genomic)

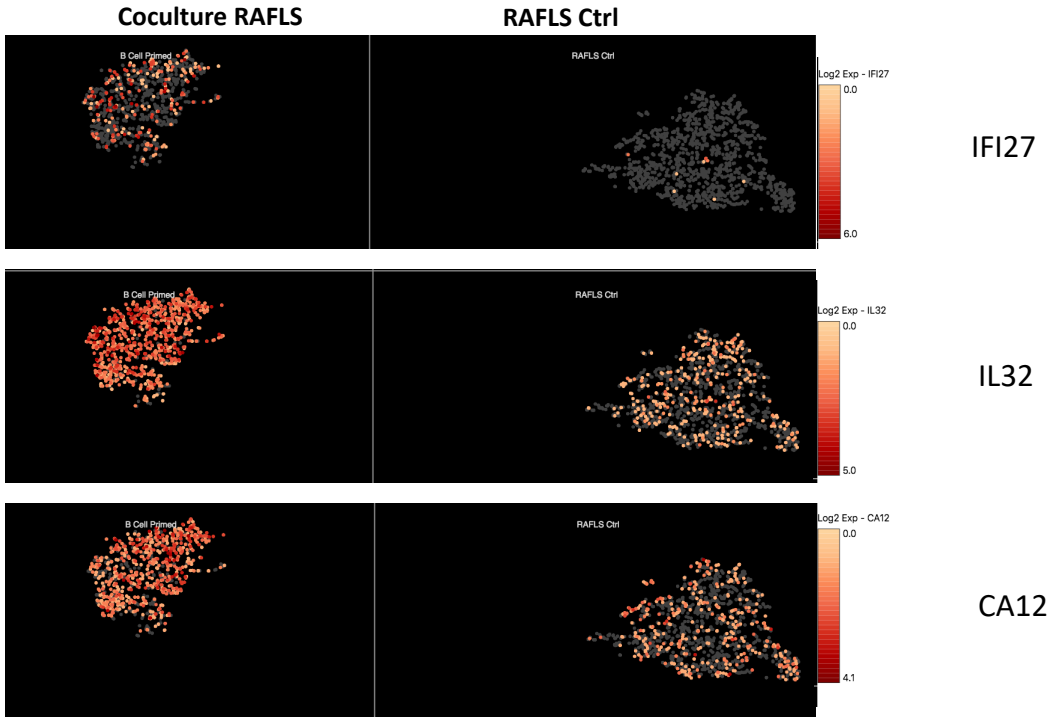


Figure 6.21. Differential gene expression between RA-FLS co-cultured with EBV+RACD19+ and control (upregulated) Results are showing that co-cultured RA-FLS (on the left) tend to upregulate expression of genes Interferon Alpha Inducible Protein 27 (IFI27), IL32, and Carbonic Anhydrase 12 (CA12). Relative expression output obtained from 10X single cells genomic sequencing. Graphs has been obtained with 10Xsoftware LOUPE (10x).

G0S2
IFIT3
XAF1
IL32
IFI6
MX1
ISG15
EPSTI1
RSPO3
SOD2
TYMP
C1QTNF1
CA12
RELB
LY6E
ABI3BP
IFITM1
STAT1
IFI35
WTAP
NFKBIA
SLC39A8
HLA-B
MSC
WARS
ICAM1
DRAM1
PSME2
SLC39A14
ATF5
CD68
PDCD5
TAPBP
TMSB4X
PLSCR1
TNFAIP6
IGFBP5
PPP1R14B
MME
UBE2L6
PRELID1
SPHK1
AGTRAP
EIF2AK2
FAM110B
RPL22L1
RND3
SERPINE1
DPYSL3
ARPC2
TNIP1
MGLL
DDIT4
OPTN
RRAS
ZNFX1
SHISA5
FTH1
ADAR
3 Cell Primed
RAFLS Ctrl

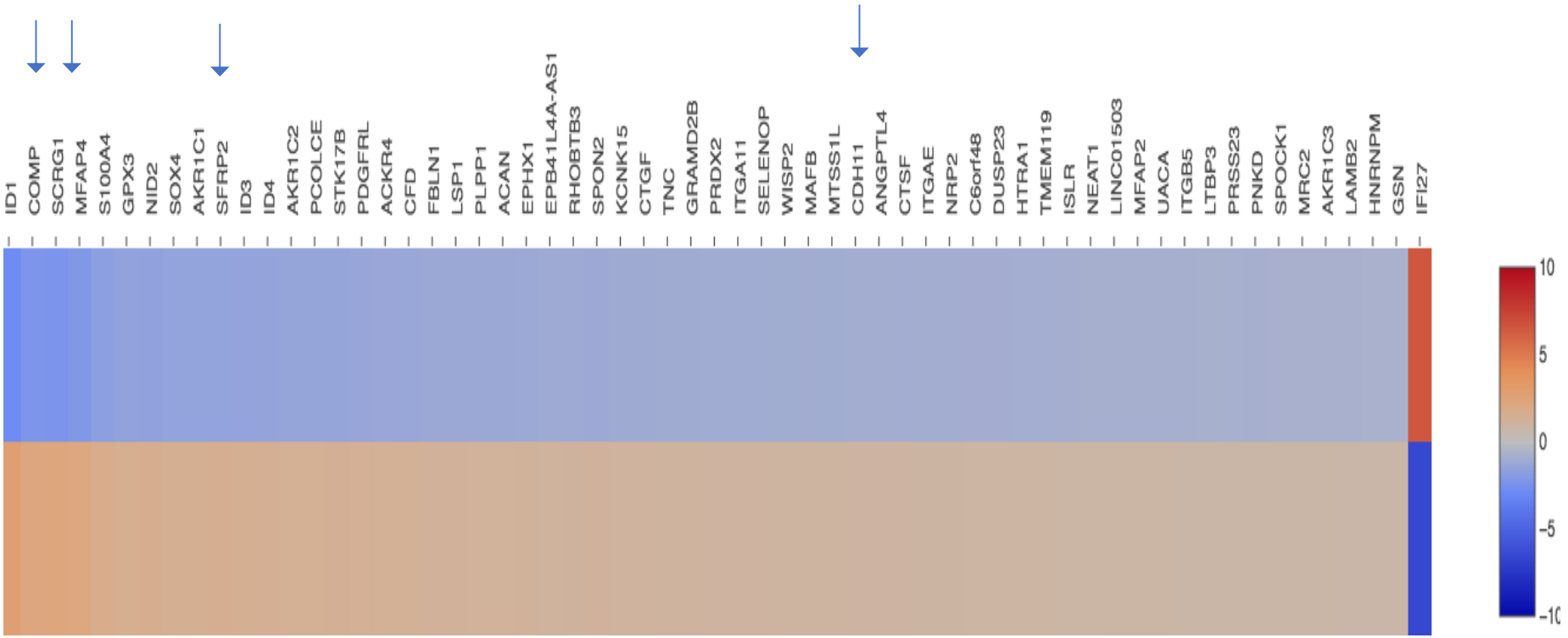
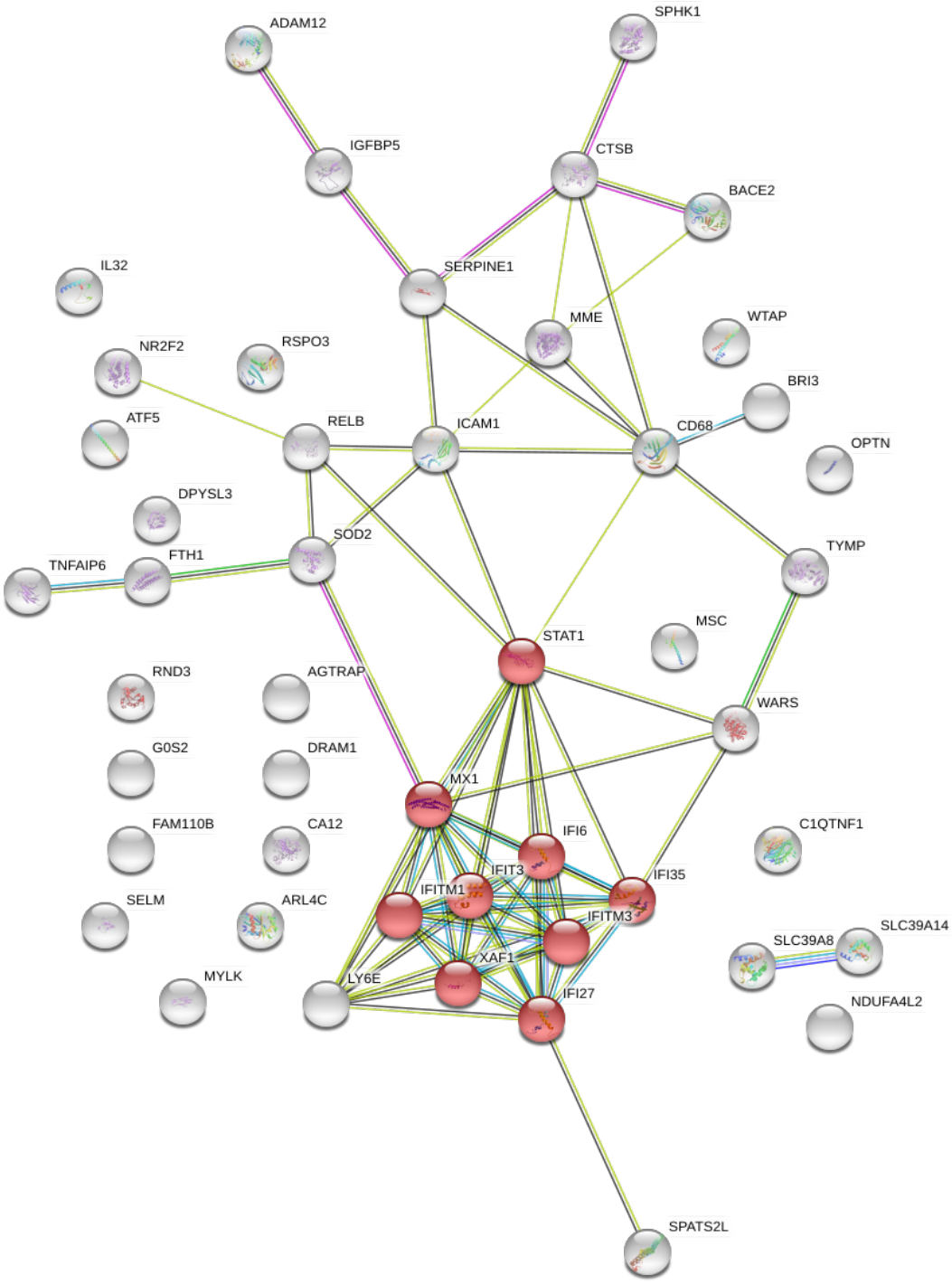
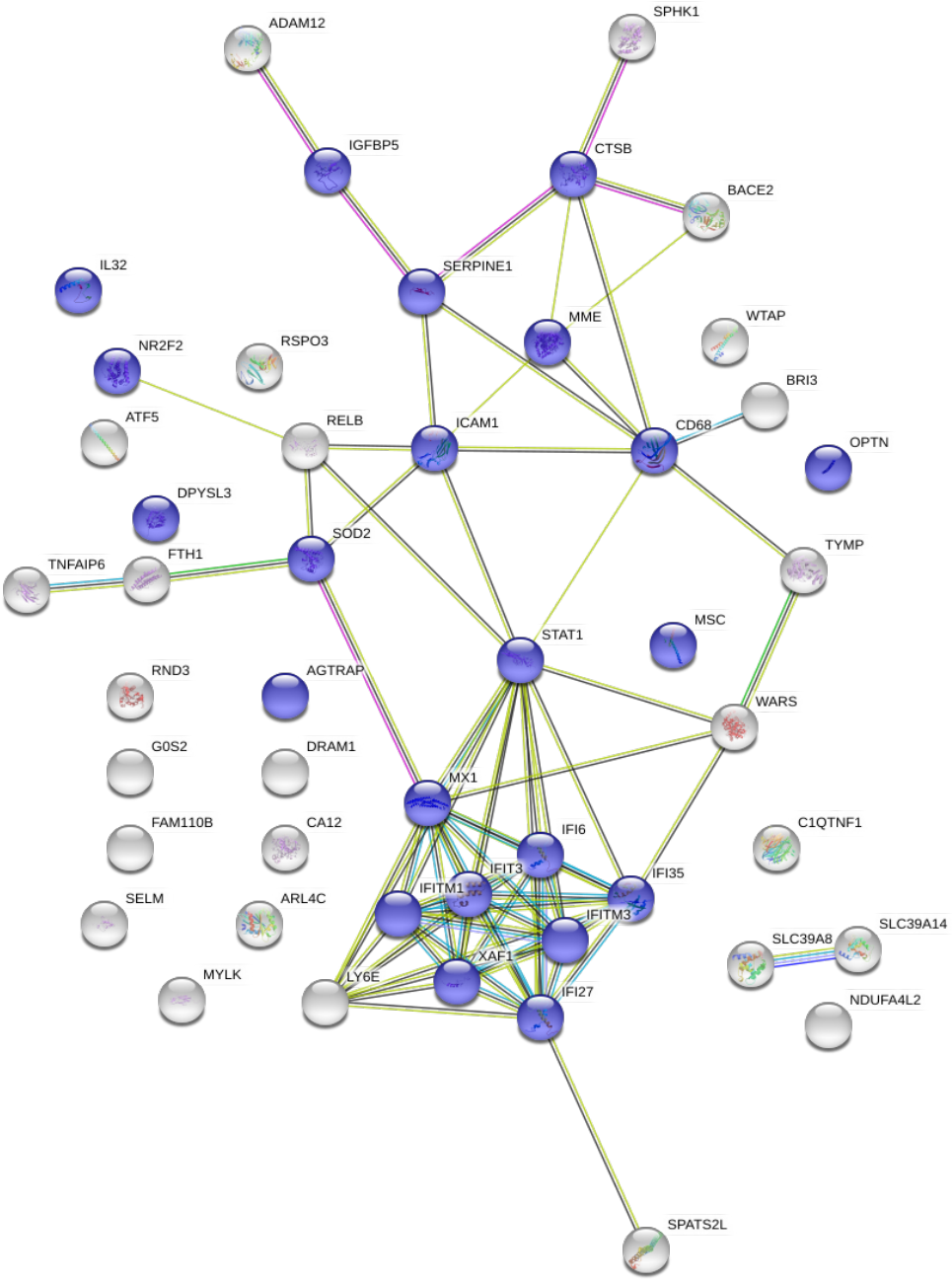


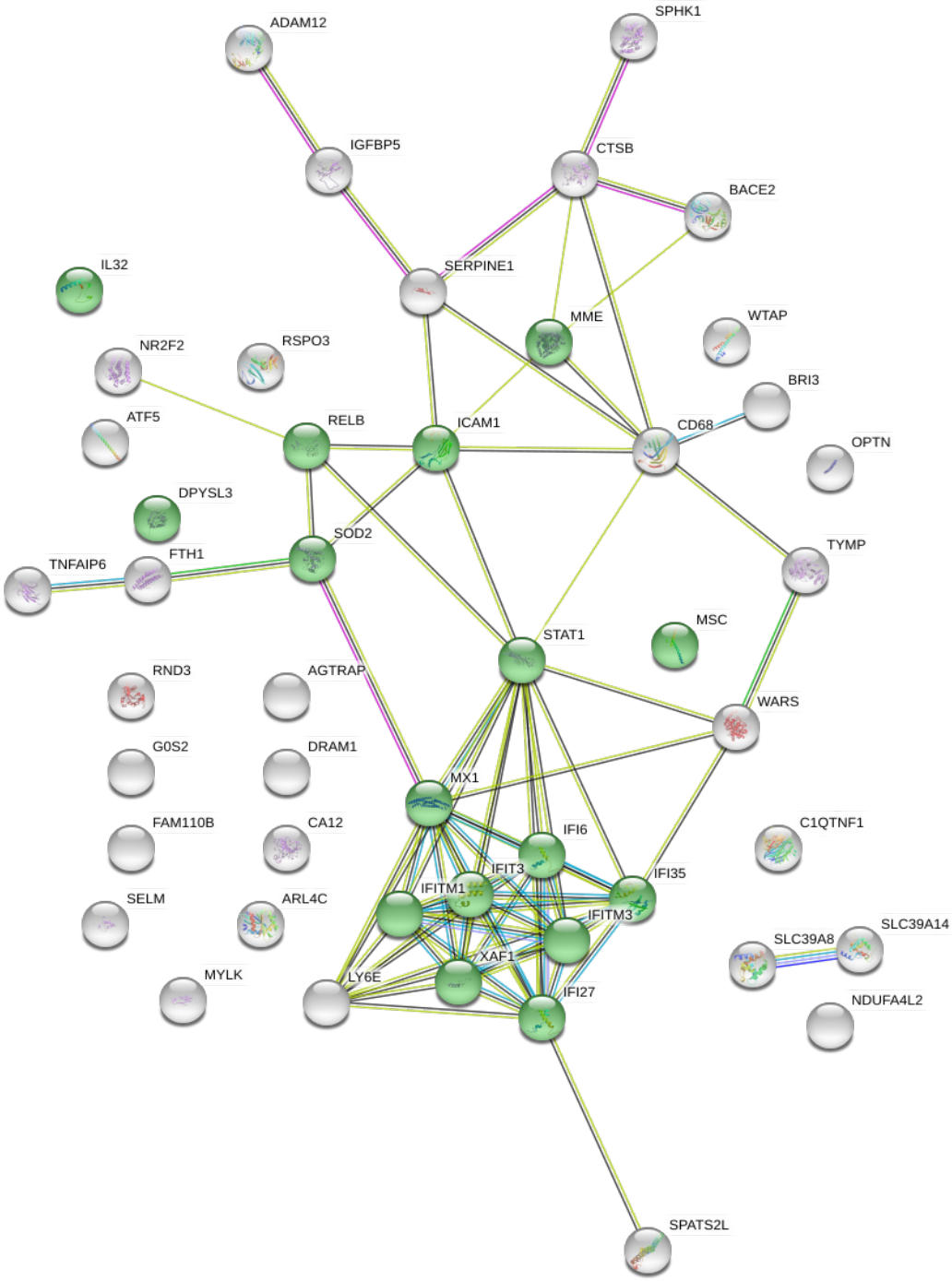
Figure 6.22. Upregulated and downregulated genes in co-culture vs control (starting from previous page). Heatmap obtained through single cells 10X genomic 5' transcriptomic analysis. B cells primed refers to RA-FLS cultured with B cells, while RA-FLS ctrl to control.



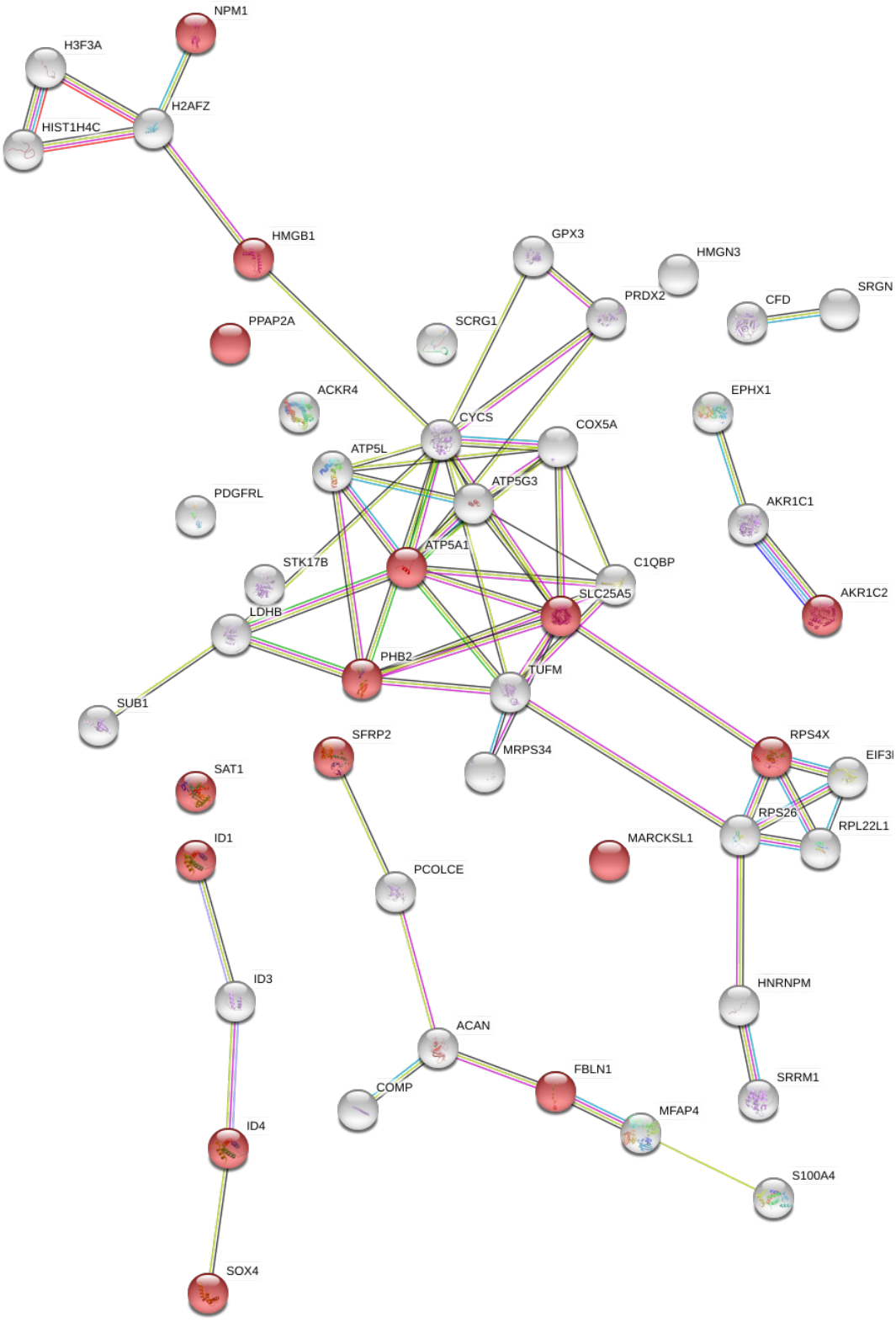
Type I interferon Signaling Pathway



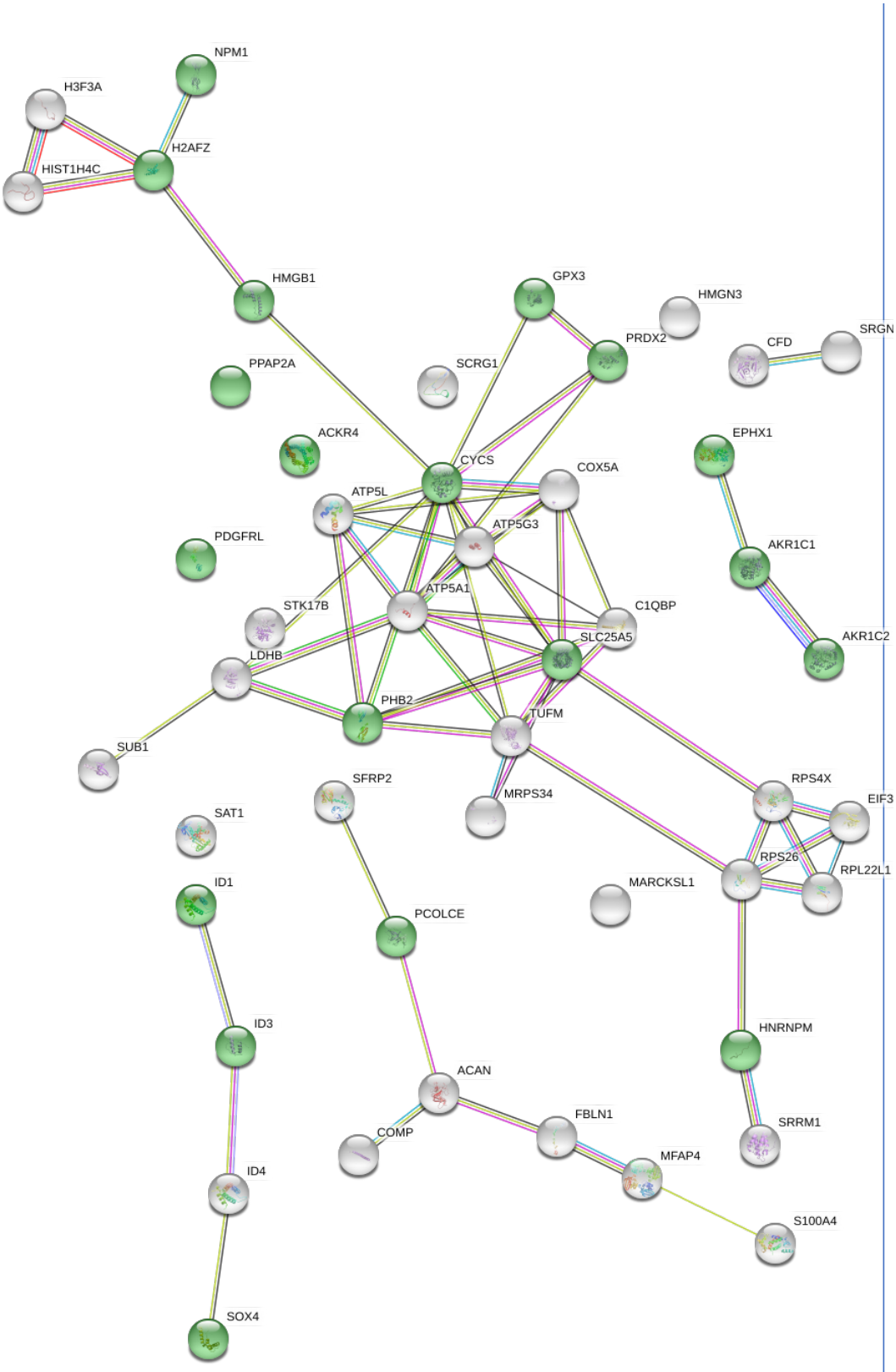
Cellular Response to organic substances



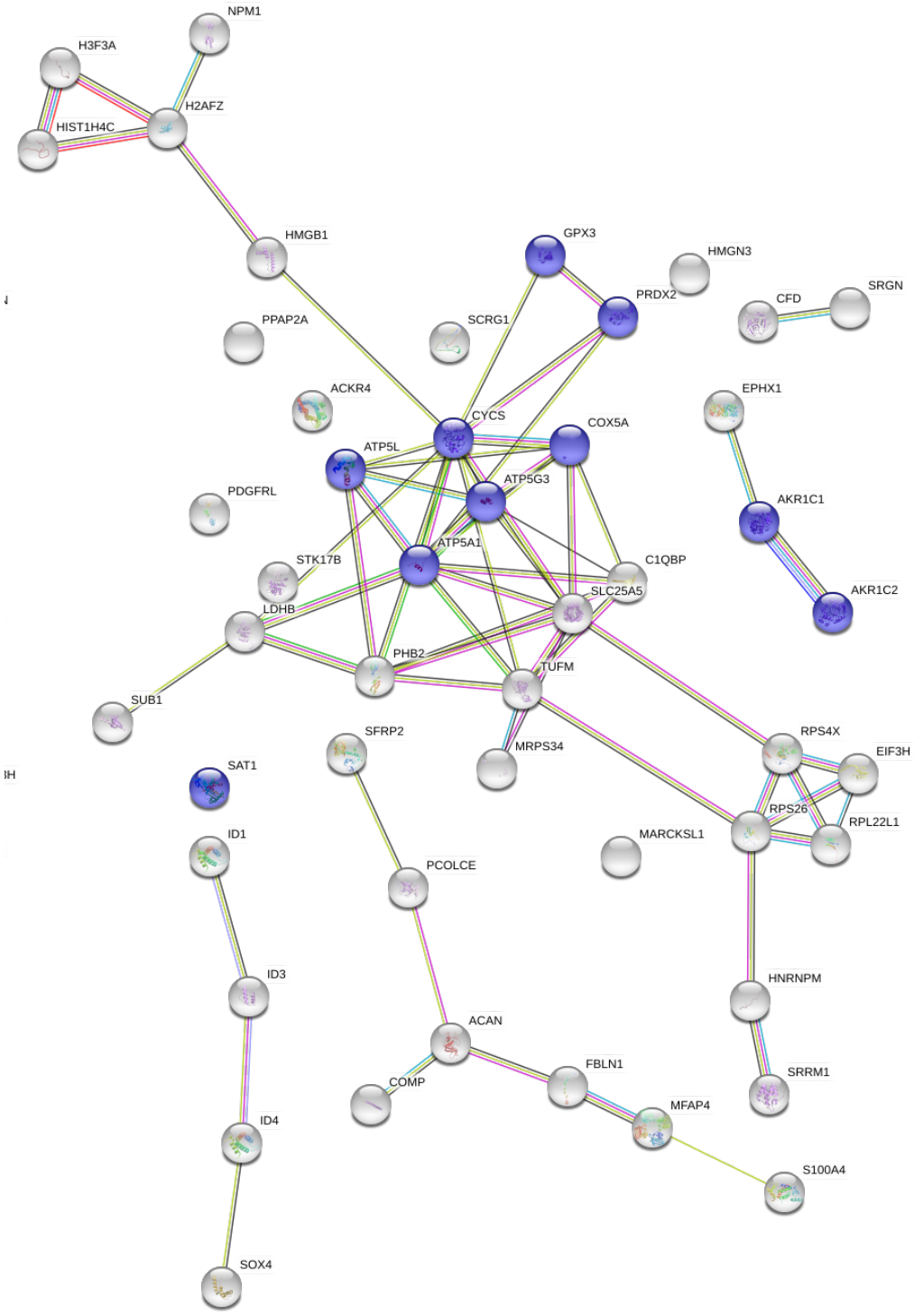
Response to Cytokine



Regulation of cell population proliferation



Drug metabolic process



Cellular response to chemical stimulus

Figure 6.23. Upregulated and downregulated genes in co-culture (previous pages): Results are showing the output obtained from STRING analysis for first 50 upregulate (Top 3) and downregulated (bottom e) genes for RA-FLS cultured with B cell, compared to RA-FLS. Output network is showing more significant interaction than expected, meaning that my proteins have more interaction among themselves than what would be expected for a random set of proteins of similar size, drawn from the genome. This enrichment indicates that the proteins are at least partially biologically connected, as a group.

Singles cells 10X genomic sequencing analysis reveal that EBV+ RACD19+ co-cultured with RA-FLS express genes important in ectopic GC formation

The EBV+RACD19+ cells cultured with the RA-FLS used in the previous experiment have been screened as well for VDJ and 5' transcriptomic. From the VDJ analysis a small group of clonotypes (mainly IgG and IgM) has emerged, (Fig.6.24). Regarding the transcriptomic analysis, the t-SNE plot obtained from the 10x analysis has divided these cells in 4 main clusters, which has given me the possibility to extrapolate an heatmap from LOUPE (Fig.6.25). From these graph it is possible to appreciate several genes important for the GC formation are, such as Epstein Barr Virus Induced 3 (EBI3), LTA and LTB (Fig.6.26). Also, several marker of proliferating B cells are upregulated, such as mki67 and also those ones connected to oncogenic viral transformation, like MYC . However, these are preliminary data and further analysis are needed.

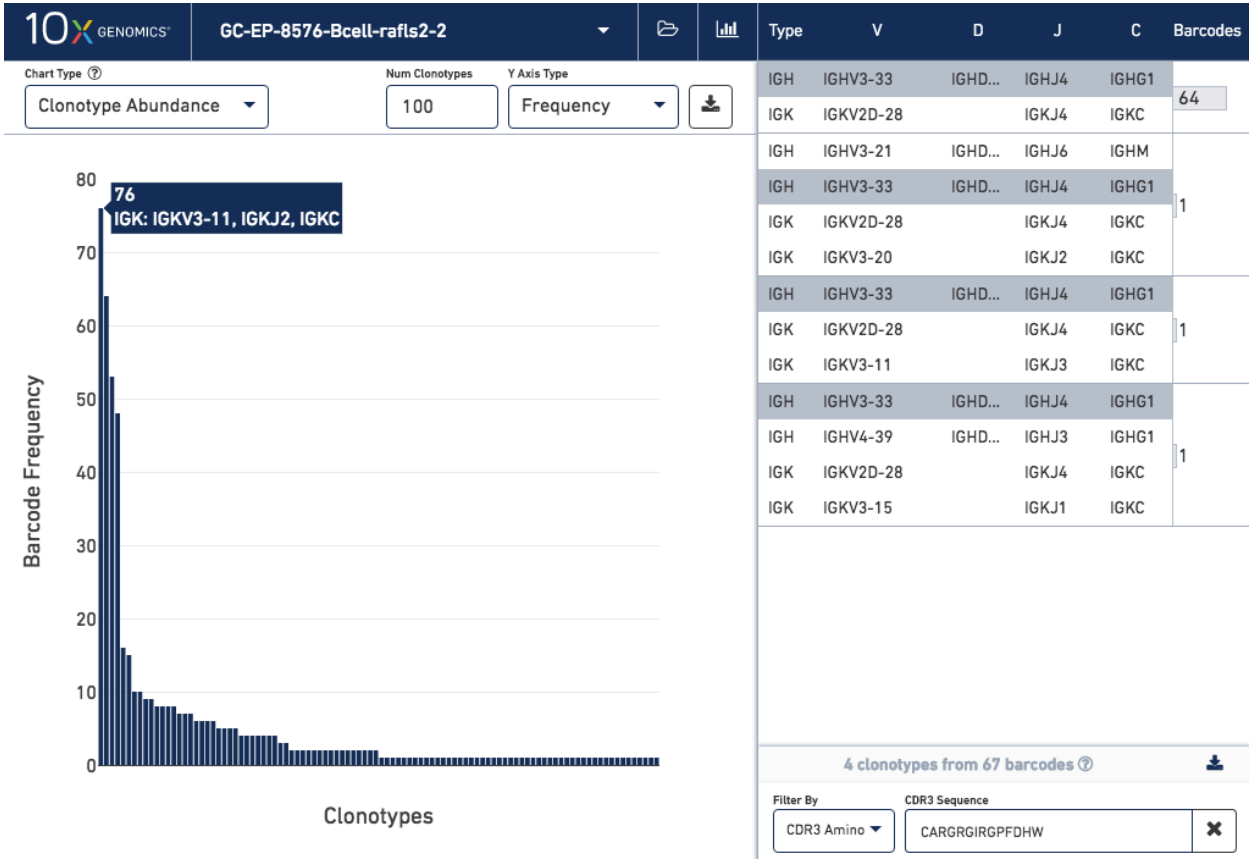


Figure 6.24. Barcode frequency from 10X genomic single cells sequencing for V(D)J: Results are showing emerging clonotypes from VDJ analysis of EBV+RACD19+ cultured with RA-FLS.

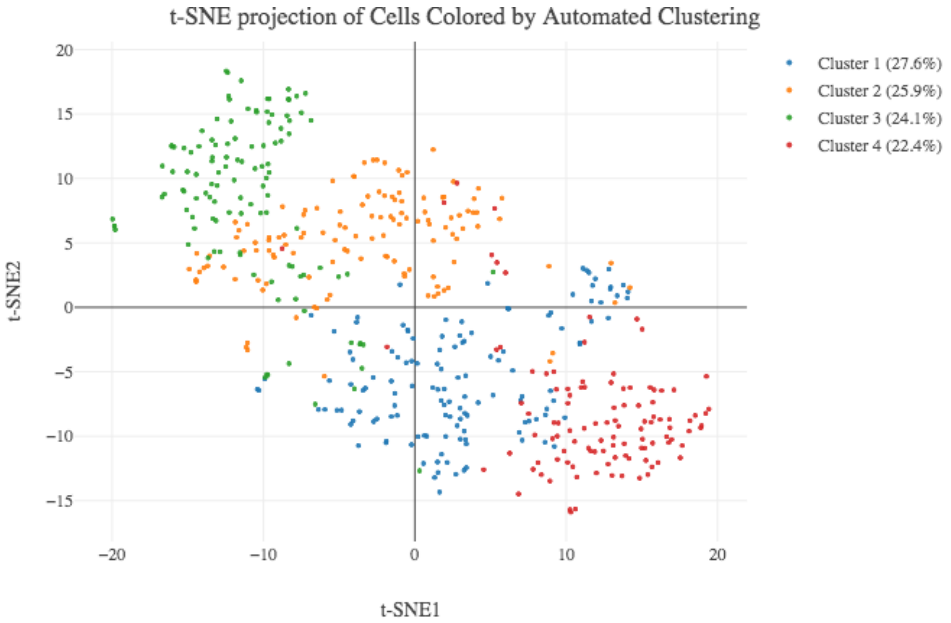


Figure 6.25. t-SNE from 10x genomic single cells sequencing for 5' transcriptomic: Results are showing 4 main clusters obtained from 10 X t-SNE analysis. From 10X genomic: "*These are the assignments of each cell-barcode to clusters by an automated clustering algorithm. The clustering groups together cells that have similar expression profiles. The axes correspond to the 2-dimensional embedding produced by the t-SNE algorithm. In this space, pairs of cells that are close to each other have more similar gene expression profiles than cells that are distant from each other*".

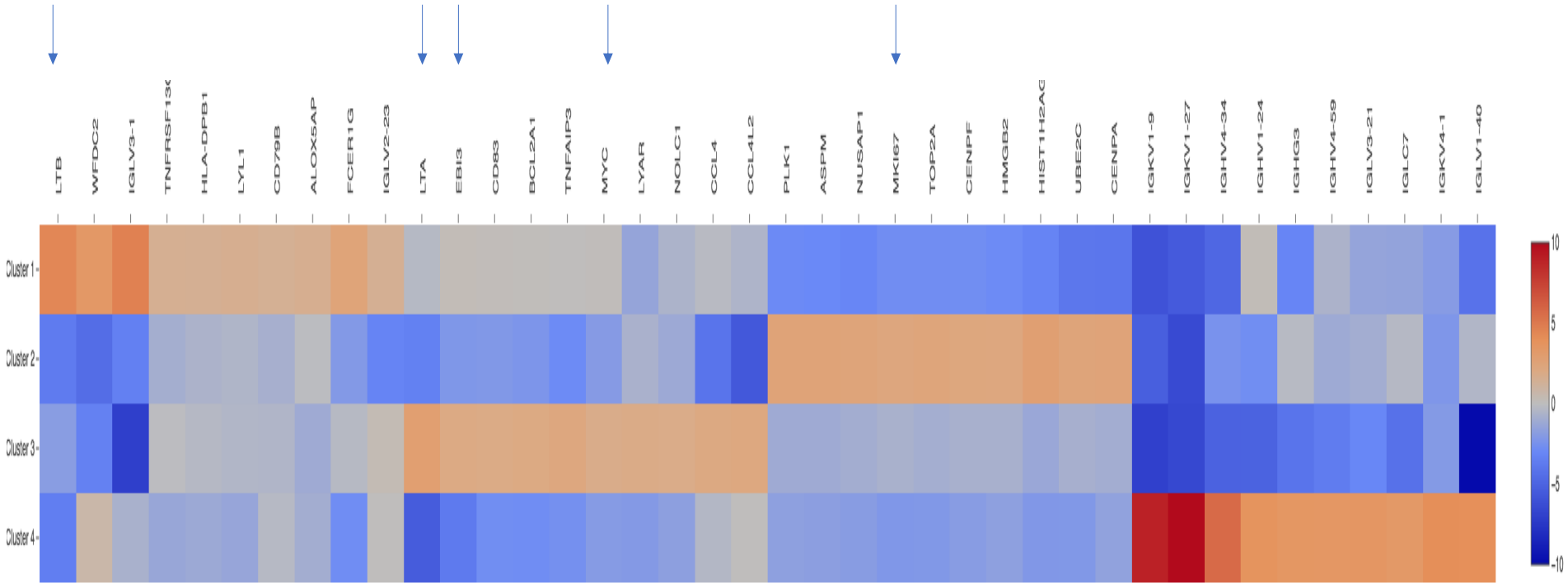


Figure 6.26. Upregulated and downregulated genes in of EBV+RACD19+ co-cultured with RA-FLS: Heatmap obtained through single cells 10X genomic 5' transcriptomic analysis. The 4 clusters obtained are showing important genes for the ectopic GC formation, like LTA, LTB and EBI3. Also, genes important in oncogenic transformation mediated by the virus are highlighted (MYC).

CHAPTER 7 DISCUSSION

Significance of EBV infection in RA as a possible etiological factor

In these PhD project I intended to clarify the intimate relationship between EBV, autoreactive B cells activation and the maintenance of autoimmunity against citrullinated antigens in the joints. I have demonstrated that synovial fluid and synovial tissue FLS enhance the long-term survival of naturally occurring EBV infected B cells obtained from circulating B cells of both RA patients and healthy donors. This work has led to the generation of one RA-derived EBV+LCL cell line (obtained with the above-mentioned feeding layer) and suggest that the co-culture of B cells with RA-FLS might be a good experimental arrangement for the establishment of LCL, which are a powerful research tool.

Also, I have investigated if - on the other hand - the co-cultured EBV infected B cells are able to enhance RA-FLS inflammatory response, which from my single-cell RNA seq data seems the case.

For last, I have tried to investigate if these cells are bearing an autoreactive phenotype, but at the moment all results are negative and further investigation is needed.

In order to obtain long-standing EBV-infected B cells, I have setup an *in vitro* co-culture system whereby RA-FLS obtained from synovial fluid and synovial tissue of ACPA+ patients were cultured with isolated CD19+ B cells from RA patients and HD controls for long-terms. RA-FLS were used as scaffolding and feeding environment for RACD19+ cells obtained from ACPA+ RA patients. CpG was added to the co-culture in order to induce EBV+ memory cells to plasma-cells differentiation for testing the presence of ACPA antibodies towards viral epitopes (total IgG, IgM and IgA; IgG-ACPA, IgG-EBNA, IgG-VCA). The latter has been chosen following Lanzavecchia and Bernasconi's protocol, where they showed that human memory B cells tend to proliferate and finally differentiate into plasmacells

in response to CpG, while Naïve B cell require simultaneously CpG and the BCR stimulation[241]. Other growth factors, cytokine and chemokine have been tested such as IL-6, IL-8, BAFF and APRIL. Furthermore, Interleukin-2 (IL-2) was added to the co-culture in order to stimulate cell survival. At the end of the co-culture single cell sorting was performed, genomic DNA was recovered to perform molecular biology analysis to detect and quantify specific EBV-related gene through the BAMHI W-repeated region in RACD19+, while RNA was isolated to test it for inflammatory related genes in RA-FLS.

In the first part of my PhD I focused on getting a better understanding on how to work with FLS from different sources (synovial fluid and synovial tissue) and to how the *in vitro* conditions can actually alter the phenotype of these cells. Apparently both fibroblast from synovial fluid and synovial tissue are behaving similarly - when cultured in the same conditions - and have been kept in culture without any sign of senescence.
I decided then to perform preliminary experiments with total PBMCs and RA-FLS from ACPA+ RA patients in order to create the best protocol for my co-culture.
Unfortunately, most of these experiments failed due to a strong immune reaction of the total PBMCs versus the RA-FLS, evidently caused by MHC mismatch detected by T cells. The heterologous nature of the assay has been an obstacle during this preliminary phase and, indeed, obtaining both FLS and B cells from the same patient has resulted challenging (and has been possible only in one case).
After performing CD19+ positive bead purification to the PBMCS prior to the co-culture, my protocol greatly improved. Also, the fact that plasmacells almost do not express CD19 has allowed to perform a sort of "autozero" in my system, regarding CSR and antibody production. At this point I have been finally able to obtain some preliminary results. The first evidences showed that RA-FLS are able to nourish B cells and keep them alive up to a maximum of 123 days and to observe the first sign of

activation and proliferation. Some of the mechanisms through which FLS helps different immune cells are well known: CD19+ cells are protected from cell death when cultured with FLS from RA through a mechanism that requires VCAM-1 and α4β1 integrin (VLA-4) and this interaction has been shown to be associated with increased expression of apoptosis inhibitors like Bcl-X[253]. Also, different cytokines such as BAFF and APRIL are produced by FLS. The former is produced in RA after engagement α5 (VLA-5) and β1 (ITGB1) integrins on the cell surface [117]. Production of these two ligands in supernatant has already been proven in my lab by Bombardieri and Kam during a short co-culture experiment (1 to 3 days). In this experiment RA-FLS – upon a TLR-3 stimulated pathways - were able to sustain tonsil naïve B cell SHM through the induction of AID expression and Ig class-switching [126]. At this point I decided to: i) standardise the protocol time and conditions; ii) increase the number of starting B cells (in order to have a greater chance of getting an EBV+ B cells); iii) test also B cells from healthy donor on the same RA-FLS. Indeed, SHM and CSR were confirmed also in my experiments through ELISA and FACS.

Specifically, FACS data have shown a significant increase of CD38+ and IgM+ subsets during co-culture in presence of EBV-driven proliferation and a smaller, but not significant, increase of switched IgG+ cells compared to control. Part of the increase in IgM+ might be due to the effect of the stimuli, since this subset is present also in the control B cells and CpG is known to induce also proliferation [241, 254]. Nevertheless, the difference between these two groups - when the cells are co-cultured - is significantly higher, so part of it is definitely induced by the RA-FLS. Interestingly, CpG does not seems to have any effect on CD38, therefore the result obtained can be genuinely connected to the presence of EBV. In fact, the increase of CD38 and IgM observed on the lymphoblast in my co-cultures matches the LCL phenotype described in literature by several authors [255-257]. Since the typical isotype of RF is IgM and since all the patient used are ACPA positive, we performed several attempts to identify

if the antibodies produced were, in fact, autoreactive. Unfortunately, most of them failed and I am speculating that the assays I have used – which are usually used to test serum – do not have enough detectability power. Also, the more sensitive VCP1 and VCP2 assay was performed in Migliorini's lab on the first batch of supernatant in 2016 and these experiments had a lower amount of B cells, thus were more diluted. Other experiments to investigate immunoreactivity have been scheduled.

I later performed a wide array of experiments to rule out the possibility that my stimuli could have any sort of effect on the RA-FLS. Neither IL-2 nor CpG had any effect on the stromal cells in my system. Namely Ospelt et al showed that TLR9 expression by FLS it is almost null [118].

Following this exclusion procedure, I have researched the possibility that EBV can infect also the RA-FLS. This should not be the case, since RA-FLS (and fibroblasts in general) do not express the surface markers necessary for viral infection, but cases of EBV identification in fibroblasts have been reported by a few authors [258]. Therefore, the sorted fibroblasts from my co-cultures experiment have been screened by qPCR for the presence of EBV and some of them resulted positive. Nevertheless, those who were positive for EBV were also positive for CD19 (data not shown). It is unlikely that this is the result of sorting errors, since the physical parameters used to sort FLS and B cells are quite different, but a scientific explanation might come from the evidences that B cell cultured for long time *in vitro* tend to actively migrate into the fibroblast cytoplasm, a fascinating mechanism called pseudoemperipolesis [259]. Hence, further investigation to understand if this event has happened in my system are ongoing, but the prospect that an EBV+ infected B cell can be internalized by RA-FLS is of great interest.

Regarding the quantification of EBV using the BamHI W repeated regions gDNA qPCR, my results showed that both RACD19+ and HdCD19+ in co-culture tend to exhibit an EBV-driven proliferation with a small significant difference between the two groups. The fact that also HdCD19+ have proliferate with the RA-FLS is not unexpected since – as above mentioned-

FLS can actively support B cells survival and anti-EBV T CD8+ specific cells are depleted from the system. It is important to mention that the results showed are an interpolation - through a standard curve - of the number of EBV copies on a given number of cells. In order to specifically quantify and determine the number of infected cells in our system we will need to use Poisson statistics, and work to generate these graphs is ongoing [260]. Moreover in 2017 Sanosyan et al have reported that reiterations of BamHI-W repeats sequence are also associated with a higher quantification variability when compared to LMP2A [261]. Nevertheless, this work was done to enhance clinical and therapeutic monitoring precision, which is beyond the scope of my project.

Additionally, proliferation of EBV+ B cells has been already observed to be increased in the presence of fibroblast, as described by Wiesner et al [262]. Indeed, cells were kept alive for an outstanding record time of 1650 days. However, these series of experiments were ran using murine fibroblastic L cells stably transfected with the human CD40 ligand and adding IL-4 to the co-culture. In my assay FLS cells are obtained from RA ACPA+ patient and neither irradiated nor modified, resembling more the pathological the patient and so are not biased as an artificial feeder cell line.

As mentioned in the results, the LCL emerged from experiment 0051 have showed an outstanding proliferative behaviour, becoming a limitless resource of naturally EBV infected B cells from an RA patient and an advantageous tool for further experiments. After the 1st co-culture, these cells have been kept alone for a week in medium, and then cultured again a 2nd time on RA-FLS without stimuli. The supernatants have been tested through LEGENDPLEX TM for several cytokines and chemokines. In this preliminary test Mip1-alfa seems to be the first molecule to be produced by the B cells despite presence of fibroblast or stimulation. It is possible that Mip1-alfa stimulates the FLS to produce high amount of IL-6 which in return can induce B cells, when infected, to proliferate. This subsequently leads to the production of CCL2 and IL8 in the first co-culture while

CXCL10, CCL5/RANTES and IL-12p70 are detected only when lymphoblasts are present. CCL5/RANTES increase might lead to the IL-8 increase detected by ELISA in the 2nd co-culture [263]. These are just speculation and further experiments are needed to understand if any of these mechanisms are related to paracrine signalling or cell to cell contact.

LCL 0051 have also been used to investigate in RA the EBV related subset reported by Megyola et al: CD58+/CD23^{hi} and CD58+/CD23^{lo} where CD23, also known as Fc epsilon RII (FcεRII) is the "low-affinity" receptor for IgE while CD58 or lymphocyte function-associated antigen 3 (LFA-3) is a cell adhesion molecule expressed on APC, (particularly macrophages) and is usually important for interaction between B and T cells[205]. My result showed that LCL 0051 have a strong presence of both CD58+/CD23^{hi} and CD58+/CD23^{lo} and that they are more likely different stages of the LCL rather that totally separated subsets with different functions. However, it is true that single cell sorted CD58+/CD23^{hi} showed higher level of the BamHI transcript EBNA-LP compared to CD58+/CD23^{lo}. Further experiments have started to investigate whether these cells are increased in autoimmune diseases (RA, SS or SLE) both in ectopic GC of synovia or salivary gland, but also at the periphery level. This is not unlikely since De Miguel et al have already reported activated CD23 hyperexpression in B cells from blood of RA patient, (but no connection to EBV were present in this work)[264]. Furthermore, these subsets might be used in order to predict those RA patients who are at risk of developing a lymphoma due to long use of immunosuppression drug, such as methotrexate or TNF blockers [64, 265].

After expanding the LCL 0051, it was clear that these cells were not losing their proliferative behaviour, so I formulated the hypothesis that they might be a possible cell line candidate. In order to understand this, I started to perform a series of screening on the LCL 0051, referred from now on as “Carejavi”. I tested these cells for the production of immunoglobulin and

the result has showed that they produce low amounts of antibodies, mainly IgM, but not ACPA, or at least not detectable.

The first evidence that these cells were all coming from a single related clonotype - IgH: VH3-7, D2-8, JH4; IgL: IGLV2-14, IGLJ2 - has been given by the VDJ analysis performed with the 10X sequencing. Intriguingly, the 5' transcriptomic analysis has showed that the Carejavi overexpress the shared epitope related gene HLA-DRB1. This is of great interest because has made me speculate that rather than an autoantibody producing cells, the main role of the Carejavi might be of an APC. This is not unlikely since Hong et al have recently demonstrate in a very elegant *in vivo* work that B cells can be the dominant APC activating CD4+ T cells in response to virus-like particles[266]. Therefore, further investigation to understand if the Carejavi are able to present autoreactive antigens to other cells are already scheduled. The possibility of having a natural APC EBV+ infected B cells obtained from an ACPA+ patient is exciting, since it can become a model for a better understanding of possible EBV-RA related mechanisms: i) binding of the MHCII by gp42, ii) investigation of disease-related protein. Also, in connection with Trier et al work, it might elucidate EBV strain-specific characteristic in ACPA+RA patients [229]. It is imperative to mention, tough, that the possibility that the Carejavi preserve the identity of the originally infected cells is almost null. The original cell would have been in latency 0/I and, once activated, these cells are not known to revert back to proliferating latency III. It is much more likely that Carejavi are just bystander LCL cells after one cycle of lytic reactivation. Furthermore, the lack of any sign of autoreactivity so far may suggest that the rare EBV+ circulating cells are absolutely not more prone to be autoreactive nor representative of the situation in the RA synovium. Indeed, Tracy et al (as already mentioned in the introduction) have demonstrated that EBV can be found in self-reactive cells but does not preferentially survive in this compartment and so the possibility that the LCL in my possess will have any autoimmune potential is extremely low. One way to address this issue would necessarily require a different

approach and an increased number of samples, as suggested by Croia in 2013: using explants from RA synovial tissue (or dissociated cultures) containing ectopic germinal centers (thus including both reactive fibroblasts and infiltrating lymphocytes) and using cyclosporin to inhibit T cell activation, could be used to immortalize the local EBV infected B cells and verify whether they indeed produce ACPA. Such an assay would be more difficult to control in term of all the different actors present in the explant (macrophages, perycites, endothelial cells, ecc ecc), but would definitely resemble more the synovial compartment *in vivo*, since the B cells would originate from the tissue and not from the peripheral blood. On a positive note, the alternative analysis performed with Seurat has shown that the Carejavi single cell dataset resemble very closely other similar LCL datasets available online: indeed, when compared to GM18502 and GM12878, the Carejavi seems to have the same mutually exclusive clustering of differentiation/immunoglobulin related genes compared to the NFkB targets. This is evident as well for CD58 and CD23/FCER2, which mostly overlap with the same cells expressing the NFkB group.

Regarding the investigation on the RA-FLS - after numerous attempts - unfortunately part of the FACS data were inconclusive. It is true that I found the best condition of detachment, preservation of surface markers and staining for FLS, but the group of markers selected in my assay to identify different fibroblast subsets (following Mizoguchi's work in 2017) have not allowed me to distinguish any possible pattern induced on the FLS by the B cells (data not shown) [149]. Also, reproducibility of a few positive experiments has been an issue. Conversely, the ELISA data on IL-6 and IL-8 were reliable. Both RA-FLS and B cell can produce IL-6, but the IL-8 identified in my co-culture seems to be produced only by the RA-FLS. I still need to elucidate whether this increase compared to control is directly related to some sort of active mechanism implying the virus or is an indirect consequence of only a higher number of cells. IL-6 is usually transiently produced in response to infective agents and trauma, but it also

involved host defence through the stimulation of acute phase reactions and immune responses. It is also involved in haematopoiesis[267].

For what concern IL-8, this specific chemokine (C-X-C motif) has many roles, like inducing chemotaxis in target cells such as neutrophils and granulocytes. Also, IL-8 is strongly correlated to angiogenesis and a connection between EBV presence, stimulation of IL-8 and neovascularization in pathological conditions has been investigated by Yoshizaki et al [268]. It is not unlikely to speculate that EBV might induce a similar mechanism in the ectopic GC of an inflamed synovia - maybe in the early phase of the disease - acting as a trigger for IL-8 production by the stromal cells compartment, finally activating some of the events which have been quite often associated with a semi-neoplastic behaviour in RA.

It is important to add that a slight increase in IL-6 production has been detected also in those samples which have not proliferated, and this sort of baseline enhancement is probably due to the feeding environment offered by the FLS.

The availability of single cells 10x genomic sequencing has pushed me to screen a co-culture experiment with this assay. These data have just been recently produced, nevertheless some of the preliminary evidence are worth to be mentioned.

The output has showed that RA-FLS co-cultured with EBV+ RACD19+ (without any stimuli) tend to switch from proliferative and invasive phenotype to a more inflammatory/cytokine producing one. Specifically, they tend to downregulate COMP, MFAP4, SCRG1 and Cadherin 11 (all genes related to the extracellular matrix and cell proliferation pathways) and to increase IFI27 or CA12 (more related to type I interferon signalling, cellular response to organic substances and response to cytokine). In particular this activation of the antiviral type 1 IFN pathway can be explained by the presence on FLS of TLR-3 and RIG-I, two PRR receptors which are known - as reported by Samanta et al - to be activated upon stimulation by EBER, which in my system is very likely to be present and have originated from the infected B cells [269]. On this topic, Stuhlmeier et

al have reported that the EBV is able to increase the production of hyaluronan on RA-FLS trough a similar mechanism [270]. Furthermore, Karpu et al have reported that triggering of the RIG-I, TLR-3 and MDA5 is able to induce the expression of CD55, but no such event was recorded in my experiment [271].

Of interest IL-32, one of the upregulated genes, is a chemokine who has been showed to prevent EBV lytic activation and has shown to have the potential to aggravate osteoclastogenesis in RA trough an IL-17 axis[272, 273].

A possible mechanism of the above-mentioned interaction is summarised in figure 8.1.

In conclusion, the relationship between EBV and RA is complex, and several mechanisms have been highlighted in this dissertation. With this work I aimed mainly to clarify the intimate relationship between EBV, autoreactive B cells activation and the maintenance of autoimmunity against citrullinated antigens in the joints. I provided evidence that synovial fluid and synovial tissue RA-FLS enhance the long-term survival of naturally occurring EBV infected B cells obtained from circulating B cells of RA patients and healthy donor and that, as a reciprocal mechanism, RA-FLS can switch phenotypes due to the interactions with the EBV+ B cells.

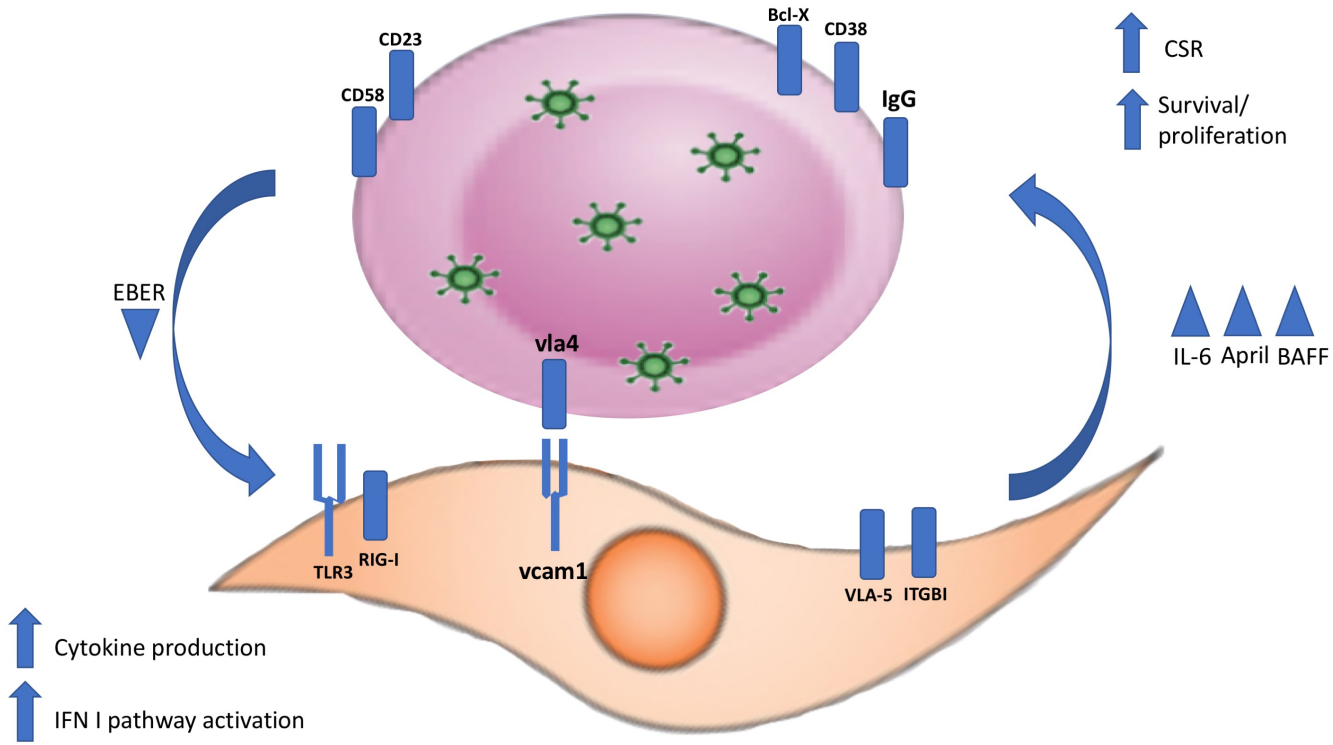


Figure 8.1. Possible mechanism of interaction between EBV+CD19+ co-cultured with RA-FLS: Scheme summarising the possible mechanism by which the interaction of RA-FLS and EBV+ infected CD19+ cells might promote RA pathology. Upon activation of VLA-5 and ITGBI, RA-FLS upregulate April and BAFF production. Also, interaction between VLA4 and VCAM1 sustain survival of B cells and increase the expression of antiapoptic signals, such as Bcl-X. This mechanism just not leads to increased survival and proliferation of B cells but also an increase of CSR. Proliferating EBV+ B cell upregulate CD23, while the production of EBER can activate the type 1 IFN pathway and the release of several cytokine, such as IL8 and IL6, which will fuel again B cell survival and inflammation.

CHAPTER 8: PLAN OF FUTURE STUDIES

Several plans for future work have been already highlighted in this manuscript. A summary is showed here:

- **i)** In order to replicates the co-culture results on a higher magnitude I will collect more RACD19+ from different donors and start a new set of co-culture experiments. Also, control co-culture will be performed using: dermal fibroblast or FLS from different RA pathotypes; ACPA- patients, EBV- patients.
- **ii)** Confirming 10X data on RA-FLS repeating the experiment with different samples.
- **iii)** Continue characterization on Carejavi cells to understand scientific value and possible potential commercial exploitation: specifically investigate the possible APC role of Carejavi regarding autoimmune antigens.
- **iv)** In case of EBV driven proliferation an extra 96well plate will be single cell sorted in order to express the monoclonal antibody and, when possible, test them for autoreactivity. Preliminary data have already shown the feasibility of this approach.
- **v)** Investigate the role of the CD23/CD58 population by: FACS and immunofluorescence in RA and other autoimmune diseases
- **vi)** Increase ELISA and Legendplex panel in order to test other molecules relevant in inflammation, such as CXCL12 or RANK-L.
- **vii)** Investigate cell to cell Interaction between B cell and FLS using immunofluorescence

APPENDIX

Flask Size	Final Volume	Seeding Density (Viable cells per flask)
T25	5-7 ml	$2.0 \times 10^5 – 6.0 \times 10^5$
T75	15 ml	$8.5 \times 10^5 – 1.3 \times 10^6$
T175	20-25 ml	$1.8 \times 10^6 – 2.4 \times 10^6$

Table 3.1. Seeding density for FLS: Note: The chart above is to be used as a guideline only. Basic seeding rule is 1.0-1.4 x 104 cells/cm^2. Readapted from Coriell Institute for Medical research, Form 1302-16 Rev H-113017.

Purity

Only LSR Fortessa 1

Marker	Fluruochrome	Filter
CD3	BV785	V 780/60
CD19	PERCP.CY5.5	B 695/40
CD20	AF488	B 530/30
CD27	APC	R 670/14
CD38	BV711	V 710/50
IgD	BV421	V 450/50
CD56	APC Cy7	R 780/60
CD14	PE	YG 582/15
LiveDead	Zombie	V 525/50

Characterization and Sorting

Both Fortessa 1 and ARIA Fusion

Marker	Fluruochrome	Filter
CD3	PE Cy5	YG 670/30
CD19	PERCP.CY5.5	B 695/40
CD20	AF488	B 530/30
CD27	APC	R 670/14
CD38	BV711	V 710/50
IgD	PE	YG 582/15
IgG	APC H7	R 780/60
IgM	BV421	V 450/50
CD24	PE/Dazzle 594	YG 610/20
LiveDead	Zombie	V 525/50

Marker	Fluruochrome	Filter
CD19	PERCP.CY5.5	B 695/40
CD20	AF488	B 530/30
CD27	APC	R 670/14
CD38	BV711	V 710/50
IgD	PE	YG 582/15
CD23	BV421	V 450/50
CD58/LFA3	PE Cy5	YG 670/30
LiveDead	Zombie	V 525/50
CD3	V500	V 525/50
CD14	BV510	V 525/50
CD56	BV510	V 525/50

Table 3.2. Purity, Characterization and Lymphoblast Panels for B Cell. Flow Cytometry Analysis panels for purity (Panel1), characterization (Panel2) and Lymphoblast (panel 3) for B cell.

Name	Cluster	Cell	Flurochrome
PTPRC	CD45	Hematopoietic lineage cells	BV510
MCAM	CD146	Pericytes	BV510
GYPA	CD253a	RedBloodCells	BV510
PECAM1	CD31	Endothelial cells	BV510
Podo		**FLS**	**AF647**
Cadherin11		**FLS**	**PE**
CD34	**CD34**	**FLS**	**BV785**
THY-1	**CD90**	**FLS**	**BV 650**
DAF	**CD55**	**FLS**	**PerCPCy5.5**
FAP		**FLS**	**AF405**
Zombie			**NIR**

Table 3.3. FLS Characterization Panel.

Gene	Forward (For) and Reverse (Rev)	Probe
W	For: AGTGGGCTTGTTTGTGACTTCA Rev: GGACTCCTGGCGCTCTGAT	TTACGTAAGCCAGACAGCAGCCAATTGTC

Table 3.4. Primer and probe for EBV BamHI W gDNA qPCR. Fl All probes are 6-FAM-TAMRA unless otherwise noted (probe 250nM/primer 900 nM)

GAPDH F	**ACAGTCCATGCCATCACTGCC**
GAPDH R	**GCCTGCTTCACCACCTTCTTG**
CD19 F	**AGAACCAGTACGGGAACGTG**
CD 19 R	**CTGCTCGGGTTTCCATAAGA**
LMP2A F	**TTCTGGCTCTTCTGGGAACA**
LMP2A R	**ATTGCCCCAATCTGAGTCCT**
EBER F	**AAAACATGCGGACCACCAGCC**
EBER R	**AGGACCTACGCTGCCCTAGA**
EBNA 1 F	**GGAGCCTGACCTGTGATCGT**
EBNA 1 R	**TAGGCCATTTCCAGGTCCTGTA**
EBNA2F	**TGGAAACCCGTCACTCTC**
EBNA2I	**TAATGGCATAGGTGGAATG**
EBNA2C	**AGGGATGCCTGGACACAAGA**
EBNA2G	**GCCTCGGTTGTGACAGAG**
EBNA2B	**TTGAAGAGTATGTCCTAAGG**

Table 3.5. List of primers for EBV PCR (Normal, Gradient, Nested)

	Patient Code		Diagnosis	Sample	Sex	Age at Collection
RA	Qmul-Rbb	057/11	RA	Synovial tissue	Male	N/A
	Qmul-Rbb	040/2016	RA	Synovial Fluid	Female	48
	Qmul-Rbb	2018/003	RA	Synovial Tissue	Male	27
	Qmul-Rbb	034/2016	RA	Synovial Fluid	Female	57
	Qmul-Rbb	023/2015	RA	Synovial Fluid	Male	N/A
	Qmul-Rbb	099/2016	RA	Peripheral Blood	Female	74
	EMR-RTB	0051	RA	Peripheral Blood	Male	46
	EMR-RTB	0060	RA	Peripheral Blood	Female	75
	EMR-RTB	0160	RA	Peripheral Blood	Male	28
	EMR-RTB	0190	RA	Peripheral Blood	Female	66
	EMR-RTB	0195	RA	Peripheral Blood	Female	32
	EMR-RTB	0022	RA	Peripheral Blood	N/A	N/A
	EMR-RTB	0023	RA	Peripheral Blood	N/A	N/A
	EMR-RTB	0029	RA	Peripheral Blood	N/A	N/A
	EMRRTB	0050	RA	Peripheral Blood	N/A	N/A
OA	Qmul-Rbb	099/2017	OA	Synovial Tissue	Female	74
	Qmul-Rbb	109/2017	OA	Synovial Tissue	Female	63

PsA	Qmul-Rbb	026/2015	Psoriatic Arthritis	Synovial Tissue	N/A	N/A
HD	EMR-RTB-HD	0040	Healthy	Peripheral Blood	Male	31
	EMR-RTB-HD	0041	Healthy	Peripheral Blood	Female	32
	EMR-RTB-HD	0047	Healthy	Peripheral Blood	Female	42
	EMR-RTB-HD	0048	Healthy	Peripheral Blood	Female	48
	EMR-RTB-HD	0049	Healthy	Peripheral Blood	Female	32
	EMR-RTB-HD	0050	Healthy	Peripheral Blood	Female	42

Table 4.1 List of Samples

Patient Code		DX	Sample	Source	Cell	Experiment			
						Experiment	Number of Experiment	Replicates (for each experiment)	Notes
Qmul-Rbb	034/2016	RA	Synovial Fluid	Knee aspiration	FLS	• Co-culture preliminary setup	• 8	• 8	• Sample have been used during set up phases for co-culture system
Qmul-Rbb	023/2015	RA	Synovial Fluid	Knee aspiration	FLS	• Co-culture preliminary setup	• 1	• 8	• Sample have been used during set up phases for co-culture system
Qmul-Rbb	099/2016	RA	Peripheral Blood	Medial cubital vein	B cells	• Co-culture preliminary setup	• 2	• 8	• Sample have been used during set up phases for co-culture system
EMR-RTB	0022	RA	Peripheral Blood	Medial cubital vein	B cells	• Co-culture preliminary setup	• 1	• 8	• Sample have been used during set up phases for co-culture system
EMR-RTB	0023	RA	Peripheral Blood	Medial cubital vein	B cells	• Co-culture preliminary setup	• 2	• 8	• Sample have been used during set up

									phases for co-culture system
EMR-RTB	0029	RA	Peripheral Blood	Medial cubital vein	B cells	• Co-culture preliminary setup	• 2	• 8	• Sample have been used during set up phases for co-culture system
EMR-RTB	0050	RA	Peripheral Blood	Medial cubital vein	B cells	• Co-culture preliminary setup	• 2	• 8	• Sample have been used during set up phases for co-culture system
Qmul-Rbb	057/11	RA	Synovial tissue	Knee Biopsy	FLS	• FACS comparison using different detaching agents	• 1	• 1 flask	• Sample used for FACS analysis
						• FACS Comparison between RA-FLS and OA-FLS at different passages	• 1	• 1 flask	• Sample used for FACS analysis
						• IL-6 and IL-8 production in RA-FLS stimulated with CpG and IL-2	• 1	• 2 wells of a 12 well plate	• SN tested on ELISA for IL-6 and IL-8
						• Relative Expression of co-culture related genes after 28 days stimulation with IL-2 and CpG on RA-FLS	• 1	• 3	• Extracted RNA has been retrotranscribed and cDNA used for qPCR

						• FACS Comparison between RA-FLS and PsA-FLS	• 1	• 1 flask	• Sample used for FACS analysis
Qmul-Rbb	040/2016	RA	Synovial Fluid	Knee aspiration	FLS	• FACS comparison using different detaching agents	• 1	• 1 flask	• Sample used for FACS analysis
						• Co-Culture	• 4	• 2 wells of a 12 well plate (30'000 FLS p/w)	• SN tested on ELISA (IL-6, IL-8, IgG, IgM, IgA) • Sample used for FACS analysis
Qmul-Rbb	2018/003	RA	Synovial Tissue	Knee Biopsy	FLS	• FACS comparison using different detaching agents	• 1	• 1 flask	• Sample used for FACS analysis
						• Co-Culture	• 6	• 2 wells of a 12 well plate (30'000 FLS p/w)	• SN tested on ELISA (IL-6, IL-8, IgG, IgM, IgA) • Sample used for FACS analysis
						• Co-culture transcriptomic analysis: reciprocal effect of RACD19+ B cells on RA-FLS	• 1	• 1 well of a 12 well plate	• Single cells RNA seq analysis on sample has been performed using 10X genomic platform

						phenotype			
Qmul-Rbb	099/2017	OA	Synovial Tissue	Joint replacement	FLS	• FACS Comparison between RA-FLS and OA-FLS at different passages	• 1	• 1 flask	• Sample used for FACS analysis
Qmul-Rbb	109/2017	OA	Synovial Tissue	Joint replacement	FLS	• FACS Comparison between RA-FLS and OA-FLS at different passages	• 1	• 1 flask	• Sample used for FACS analysis
Qmul-Rbb	026/2015	PsA	Synovial Tissue	Knee Biopsy	FLS	• FACS Comparison between RA-FLS and PsA-FLS	• 1	• 1 flask	• Sample used for FACS analysis
EMR-RTB	0051	RA	Peripheral Blood	Medial cubital vein	CD19+	• Co-Culture	• 1	• 2 wells of a 12 well plate (360'000 CD19+ cells por well)	• SN tested on ELISA (IL-6, IL-8, IgG, IgM, IgA) • Sample used for FACS analysis • LCL obtained from this sample (Carejavi) • Multiplex Panel • BAMHIW EBV qPCR on gDNA
						• Expression of CD58 and	• 1	• 2 wells of a 48	• Sample used for FACS

						CD23 marker by FACS in experiment 0051		wells plate	analysis
						• EBV EBNA-LP mRNA level comparison in $CD58^{+}/CD23^{hi}$ $CD58^{+}/CD23l^{ow}$ cells when analysed at single cell level	• 1	• 48 wells of a 96 wells plate for each condition (1 cell each well)	• Extracted RNA has been retrotranscribed and cDNA used for qPCR
						• EBV Copies from separation of subsets $CD58^{+}/CD23^{hi}$ and $CD58+/CD23^{low}$	• 1	• 2 wells of a 12 well plate (360'000 CD19+ cells por well)	• Cells have been sorted according to their expression of markers and cultured for 14 days • BAMHIW EBV qPCR on gDNA
						• Singles cells 10X genomics chromium sequencing and STRING analysis on lymphoblasts 0051 (VDJ and 5' RNA seq)	• 1	• 1 well of a 12 well plate	• Single cells RNA seq analysis on sample has been performed using 10X genomic platform
EMR-RTB	0060	RA	Peripheral Blood	Medial cubital vein	CD19+	• Co-Culture	• 1	• 2 wells of a 12 well plate (360'000 CD19+	• SN tested on ELISA (IL-6, IL-8, IgG, IgM, IgA) • Sample used for FACS analysis

								cells por well)	• BAMHIW EBV qPCR on gDNA
EMR-RTB	0160	RA	Peripheral Blood	Medial cubital vein	CD19+	• Co-Culture	• 1	• 2 wells of a 12 well plate (360'000 CD19+ cells por well)	• SN tested on ELISA (IL-6, IL-8, IgG, IgM, IgA) • Sample used for FACS analysis • BAMHIW EBV qPCR on gDNA
EMR-RTB	0190	RA	Peripheral Blood	Medial cubital vein	CD19+	• Co-Culture	• 1	• 2 wells of a 12 well plate (360'000 CD19+ cells por well)	• SN tested on ELISA (IL-6, IL-8, IgG, IgM, IgA) • Sample used for FACS analysis • BAMHIW EBV qPCR on gDNA
EMR-RTB	0195	RA	Peripheral Blood	Medial cubital vein	CD19+	• Co-Culture	• 1	• 2 wells of a 12 well plate (360'000 CD19+ cells por well)	• SN tested on ELISA (IL-6, IL-8, IgG, IgM, IgA) • Sample used for FACS analysis • BAMHIW EBV qPCR on gDNA
						• Co-culture transcriptomic	• 1	• 1 well of a 12 well	• Single cells RNA seq analysis on sample

						analysis: reciprocal effect of RACD19+ B cells on RA-FLS phenotype		plate	has been performed using 10X genomic platform
EMR-RTB-HD	0040	HD	Peripheral Blood	Medial cubital vein	CD19+	• Co-Culture	• 1	• 2 wells of a 12 well plate (360'000 CD19+ cells por well)	• SN tested on ELISA (IL-6, IL-8, IgG, IgM, IgA) • Sample used for FACS analysis • BAMHIW EBV qPCR on gDNA
EMR-RTB-HD	0041	HD	Peripheral Blood	Medial cubital vein	CD19+	• All preliminary co-culture set up	• 3	• 1 condition	• Sample have been used to assess purity and antigen masking in CD19+ cells
EMR-RTB-HD	0047	HD	Peripheral Blood	Medial cubital vein	CD19+	• Co-Culture	• 1	• 2 wells of a 12 well plate (360'000 CD19+ cells por well)	• SN tested on ELISA (IL-6, IL-8, IgG, IgM, IgA) • Sample used for FACS analysis • BAMHIW EBV qPCR on gDNA
EMR-	0048	HD	Peripher	Medial	CD19	• Co-Culture	• 1	• 2 wells of	• SN tested on ELISA

RTB-HD			al Blood	cubital vein	+			a 12 well plate (360'000 CD19+ cells por well)	(IL-6, IL-8, IgG, IgM, IgA) • Sample used for FACS analysis • BAMHIW EBV qPCR on gDNA
EMR-RTB-HD	0049	HD	Peripheral Blood	Medial cubital vein	CD19+	• Co-Culture	• 1	• 2 wells of a 12 well plate (360'000 CD19+ cells por well)	• SN tested on ELISA (IL-6, IL-8, IgG, IgM, IgA) • Sample used for FACS analysis • BAMHIW EBV qPCR on gDNA
EMR-RTB-HD	0050	HD	Peripheral Blood	Medial cubital vein	CD19+	• Co-Culture	• 1	• 2 wells of a 12 well plate (360'000 CD19+ cells por well)	• SN tested on ELISA (IL-6, IL-8, IgG, IgM, IgA) • Sample used for FACS analysis • BAMHIW EBV qPCR on gDNA

Table 4.2 List of Samples according to experiment

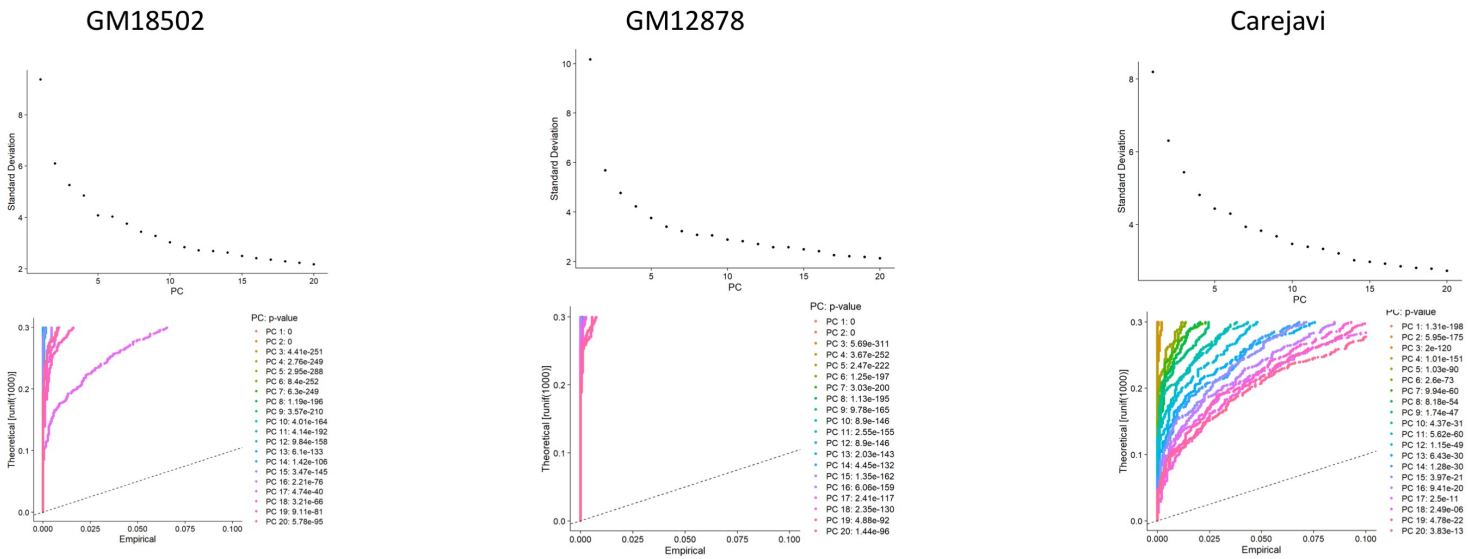
Marker	Flurochrome	Clone
CD3	FITC	HIT3a
CD14	PE	MΦP9
CD27	APC	MT271
CD38	By711	HIT2
IgD	BV421	IA6-2
CD19	PerCP.Cy5.5	SJ25C1

Marker	Flurochrome	Clone
CD3	BV785	OKT3
CD14	PE	MΦP9
CD27	APC	MT271
CD56	APC Cy7	HCD56
IgD	BV421	IA6-2
CD19	PerCP.Cy5.5	SJ25C1
CD20	AF488	2H7

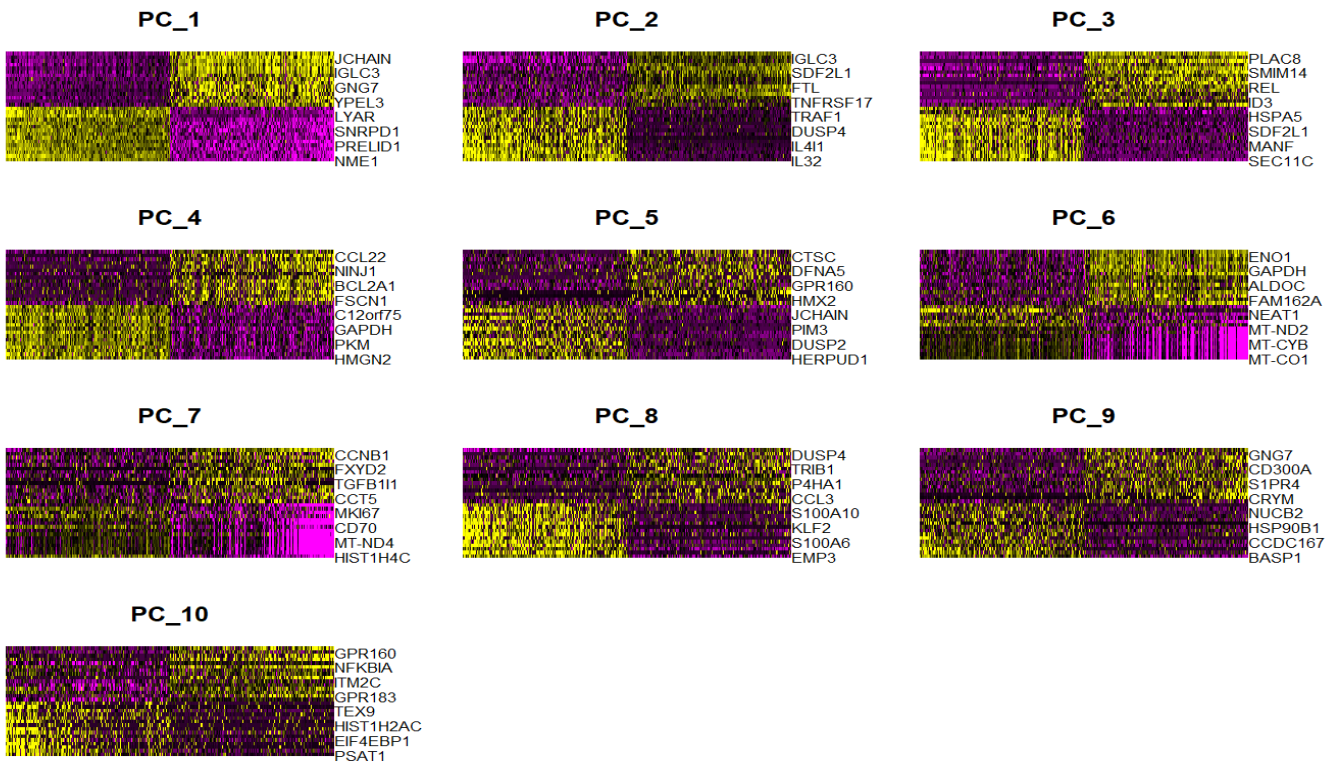
Table 4.3. 2nd and 3rd Facs panel for purity of CD19+positive selection.

A4.TSLP
A5.CXCL10
A6.TNFa
A7.IL-6
A8.CCL2
A10.CCL5
B2.RANKL
B3.MIP-1a
B4.IL-12p70
B5.IL-27
B6.APRIL
B7.BAFF
B9.IL-22

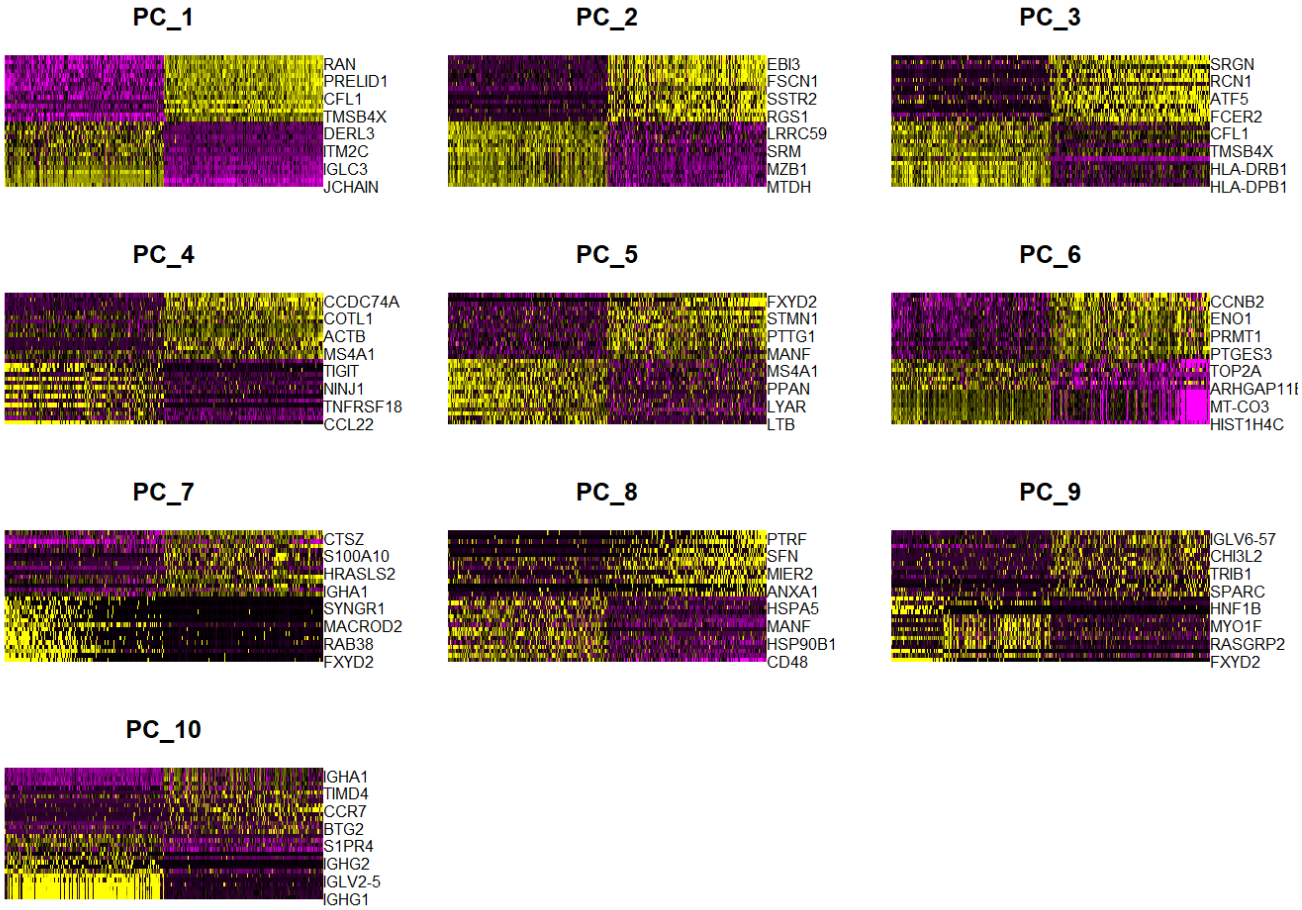
Table 6.1. Multiplex Panel.



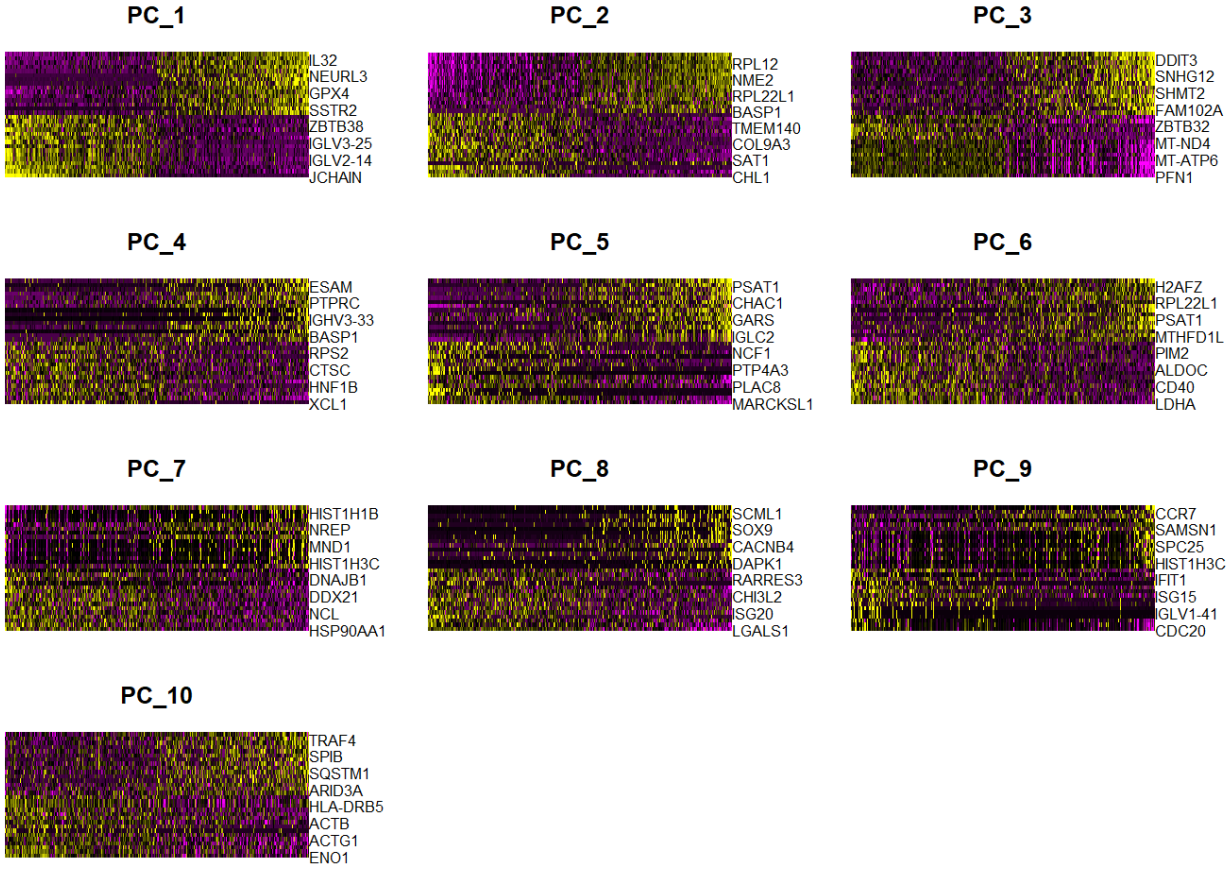
Supplementary Figure 6.1. Elbow plot and Jackstraw plot for Carejavi, GM18502, GM12878. Graph are showing Elbow plot (a ranking of principle components based on the percentage of variance explained by each one) and Jackstraw plot (visualization tool for comparing the distribution of p-values for each PC with a uniform distribution (dashed line; significant PCs will show a strong enrichment of features with low p-values (solid curve above the dashed line)).



Supplementary Figure 6.2. Dimensional Heatmap of the first 10 PCs for GM18502 (4 genes).



Supplementary Figure 6.3. Dimensional Heatmap of the first 10 PCs for GM12878 (4 genes).



Supplementary Figure 6.4. Dimensional Heatmap of the first 10 PCs for Carejavi (4 genes).

REFERENCES

1. Strand, V., R. Kimberly, and J.D. Isaacs, *Biologic therapies in rheumatology: lessons learned, future directions.* Nat Rev Drug Discov, 2007. **6**(1): p. 75-92.
2. Wasserman, A.M., *Diagnosis and management of rheumatoid arthritis.* Am Fam Physician, 2011. **84**(11): p. 1245-52.
3. Combe, B., et al., *EULAR recommendations for the management of early arthritis: report of a task force of the European Standing Committee for International Clinical Studies Including Therapeutics (ESCISIT).* Ann Rheum Dis, 2007. **66**(1): p. 34-45.
4. Aletaha, D., et al., *2010 Rheumatoid arthritis classification criteria: an American College of Rheumatology/European League Against Rheumatism collaborative initiative.* Arthritis Rheum, 2010. **62**(9): p. 2569-81.
5. Mochan, E. and M.H. Ebell, *Predicting rheumatoid arthritis risk in adults with undifferentiated arthritis.* Am Fam Physician, 2008. **77**(10): p. 1451-3.
6. van der Helm-van Mil, A.H., et al., *A prediction rule for disease outcome in patients with recent-onset undifferentiated arthritis: how to guide individual treatment decisions.* Arthritis Rheum, 2007. **56**(2): p. 433-40.
7. England, B.R., et al., *2019 Update of the American College of Rheumatology Recommended Rheumatoid Arthritis Disease Activity Measures.* Arthritis Care Res (Hoboken), 2019. **71**(12): p. 1540-1555.
8. Silman, A.J. and J.E. Pearson, *Epidemiology and genetics of rheumatoid arthritis.* Arthritis Res, 2002. **4 Suppl 3**: p. S265-72.
9. Del Puente, A., et al., *High incidence and prevalence of rheumatoid arthritis in Pima Indians.* Am J Epidemiol, 1989. **129**(6): p. 1170-8.
10. MacGregor, A.J., et al., *Characterizing the quantitative genetic contribution to rheumatoid arthritis using data from twins.* Arthritis Rheum, 2000. **43**(1): p. 30-7.
11. Seldin, M.F., et al., *The genetics revolution and the assault on rheumatoid arthritis.* Arthritis Rheum, 1999. **42**(6): p. 1071-9.
12. Gregersen, P.K., J. Silver, and R.J. Winchester, *The shared epitope hypothesis. An approach to understanding the molecular genetics of susceptibility to rheumatoid arthritis.* Arthritis Rheum, 1987. **30**(11): p. 1205-13.
13. Wucherpfennig, K.W. and J.L. Strominger, *Selective binding of self peptides to disease-associated major histocompatibility complex (MHC) molecules: a mechanism for MHC-linked susceptibility to human autoimmune diseases.* J Exp Med, 1995. **181**(5): p. 1597-601.

14. Bhayani, H.R. and S.M. Hedrick, *The role of polymorphic amino acids of the MHC molecule in the selection of the T cell repertoire.* J Immunol, 1991. **146**(4): p. 1093-8.
15. Huizinga, T.W., et al., *Refining the complex rheumatoid arthritis phenotype based on specificity of the HLA-DRB1 shared epitope for antibodies to citrullinated proteins.* Arthritis Rheum, 2005. **52**(11): p. 3433-8.
16. Zanelli, E., et al., *An extended HLA-DQ-DR haplotype rather than DRB1 alone contributes to RA predisposition.* Immunogenetics, 1998. **48**(6): p. 394-401.
17. du Montcel, S.T., et al., *New classification of HLA-DRB1 alleles supports the shared epitope hypothesis of rheumatoid arthritis susceptibility.* Arthritis Rheum, 2005. **52**(4): p. 1063-8.
18. Miyadera, H. and K. Tokunaga, *Associations of human leukocyte antigens with autoimmune diseases: challenges in identifying the mechanism.* J Hum Genet, 2015. **60**(11): p. 697-702.
19. Guthrie, K.A., N.R. Tishkevich, and J.L. Nelson, *Non-inherited maternal human leukocyte antigen alleles in susceptibility to familial rheumatoid arthritis.* Ann Rheum Dis, 2009. **68**(1): p. 107-9.
20. Molitor, M.L. and W.J. Burlingham, *Immunobiology of exposure to non-inherited maternal antigens.* Front Biosci, 2007. **12**: p. 3302-11.
21. Klareskog, L., et al., *Immunity to citrullinated proteins in rheumatoid arthritis.* Annu Rev Immunol, 2008. **26**: p. 651-75.
22. Krishnamurthy, A., et al., *Identification of a novel chemokine-dependent molecular mechanism underlying rheumatoid arthritis-associated autoantibody-mediated bone loss.* Ann Rheum Dis, 2016. **75**(4): p. 721-9.
23. Snir, O., et al., *Multiple antibody reactivities to citrullinated antigens in sera from patients with rheumatoid arthritis: association with HLA-DRB1 alleles.* Ann Rheum Dis, 2009. **68**(5): p. 736-43.
24. Yazbek, M.A., et al., *Association analysis of anti-Epstein-Barr nuclear antigen-1 antibodies, anti-cyclic citrullinated peptide antibodies, the shared epitope and smoking status in Brazilian patients with rheumatoid arthritis.* Clinics (Sao Paulo), 2011. **66**(8): p. 1401-6.
25. Orozco, G. and A. Barton, *Update on the genetic risk factors for rheumatoid arthritis.* Expert Rev Clin Immunol, 2010. **6**(1): p. 61-75.
26. Raychaudhuri, S., et al., *Common variants at CD40 and other loci confer risk of rheumatoid arthritis.* Nat Genet, 2008. **40**(10): p. 1216-23.
27. Arleevskaya, M.I., et al., *Editorial: Microbial and Environmental Factors in Autoimmune and Inflammatory Diseases.* Front Immunol, 2017. **8**: p. 243.
28. Anaya, J.M., et al., *The Autoimmune Ecology.* Front Immunol, 2016. **7**: p. 139.
29. Hurst, J. and P. von Landenberg, *Toll-like receptors and autoimmunity.* Autoimmun Rev, 2008. **7**(3): p. 204-8.

30. Wahamaa, H., et al., *High mobility group box protein 1 in complex with lipopolysaccharide or IL-1 promotes an increased inflammatory phenotype in synovial fibroblasts.* Arthritis Res Ther, 2011. **13**(4): p. R136.
31. Schrijver, I.A., et al., *Antigen-presenting cells containing bacterial peptidoglycan in synovial tissues of rheumatoid arthritis patients coexpress costimulatory molecules and cytokines.* Arthritis Rheum, 2000. **43**(10): p. 2160-8.
32. Richez, C., et al., *Role for toll-like receptors in autoimmune disease: the example of systemic lupus erythematosus.* Joint Bone Spine, 2011. **78**(2): p. 124-30.
33. Harris, D.P., et al., *Reciprocal regulation of polarized cytokine production by effector B and T cells.* Nat Immunol, 2000. **1**(6): p. 475-82.
34. Strickland, F.M., et al., *Environmental exposure, estrogen and two X chromosomes are required for disease development in an epigenetic model of lupus.* J Autoimmun, 2012. **38**(2-3): p. J135-43.
35. Rojas-Villarraga, A., J.V. Torres-Gonzalez, and A.M. Ruiz-Sternberg, *Safety of hormonal replacement therapy and oral contraceptives in systemic lupus erythematosus: a systematic review and meta-analysis.* PLoS One, 2014. **9**(8): p. e104303.
36. Molina, R., et al., *Primary Sjogren's syndrome in men. Clinical, serologic, and immunogenetic features.* Am J Med, 1986. **80**(1): p. 23-31.
37. Oliver, J.E. and A.J. Silman, *Why are women predisposed to autoimmune rheumatic diseases?* Arthritis Res Ther, 2009. **11**(5): p. 252.
38. Mustafa, A., et al., *Gestational exposure to 2,3,7,8-tetrachlorodibenzo-p-dioxin disrupts B-cell lymphopoiesis and exacerbates autoimmune disease in 24-week-old SNF1 mice.* Toxicol Sci, 2009. **112**(1): p. 133-43.
39. Blander, J.M., M.B. Torchinsky, and L. Campisi, *Revisiting the old link between infection and autoimmune disease with commensals and T helper 17 cells.* Immunol Res, 2012. **54**(1-3): p. 50-68.
40. Somers, E.C. and B.C. Richardson, *Environmental exposures, epigenetic changes and the risk of lupus.* Lupus, 2014. **23**(6): p. 568-76.
41. Costenbader, K.H., et al., *Cigarette smoking and the risk of systemic lupus erythematosus: a meta-analysis.* Arthritis Rheum, 2004. **50**(3): p. 849-57.
42. Manfredsdottir, V.F., et al., *The effects of tobacco smoking and rheumatoid factor seropositivity on disease activity and joint damage in early rheumatoid arthritis.* Rheumatology (Oxford), 2006. **45**(6): p. 734-40.
43. Pedersen, M., et al., *Strong combined gene-environment effects in anti-cyclic citrullinated peptide-positive rheumatoid arthritis: a nationwide case-control study in Denmark.* Arthritis Rheum, 2007. **56**(5): p. 1446-53.

44. Padyukov, L., et al., *A gene-environment interaction between smoking and shared epitope genes in HLA-DR provides a high risk of seropositive rheumatoid arthritis.* Arthritis Rheum, 2004. **50**(10): p. 3085-92.
45. Damgaard, D., et al., *Smoking is associated with increased levels of extracellular peptidylarginine deiminase 2 (PAD2) in the lungs.* Clin Exp Rheumatol, 2015. **33**(3): p. 405-8.
46. Karlson, E.W. and K. Deane, *Environmental and gene-environment interactions and risk of rheumatoid arthritis.* Rheum Dis Clin North Am, 2012. **38**(2): p. 405-26.
47. Lee, Y.H., S.C. Bae, and G.G. Song, *Coffee or tea consumption and the risk of rheumatoid arthritis: a meta-analysis.* Clin Rheumatol, 2014. **33**(11): p. 1575-83.
48. Horwitz, H., B. Ahlgren, and E. Naerum, *Effect of occupation on risk of developing MS: an insurance cohort study.* BMJ Open, 2013. **3**(6).
49. Marie, I., et al., *Prospective study to evaluate the association between systemic sclerosis and occupational exposure and review of the literature.* Autoimmun Rev, 2014. **13**(2): p. 151-6.
50. Minden, S., et al., *Disease-modifying agents in the Sonya Slifka Longitudinal Multiple Sclerosis Study.* Mult Scler, 2008. **14**(5): p. 640-55.
51. Benedek, T.G., *The history of bacteriologic concepts of rheumatic fever and rheumatoid arthritis.* Semin Arthritis Rheum, 2006. **36**(2): p. 109-23.
52. Sandberg, M.E., et al., *Recent infections are associated with decreased risk of rheumatoid arthritis: a population-based case-control study.* Ann Rheum Dis, 2015. **74**(5): p. 904-7.
53. Leirisalo-Repo, M., *Early arthritis and infection.* Curr Opin Rheumatol, 2005. **17**(4): p. 433-9.
54. Arleevskaya, M.I., et al., *How Rheumatoid Arthritis Can Result from Provocation of the Immune System by Microorganisms and Viruses.* Front Microbiol, 2016. **7**: p. 1296.
55. Perricone, C., et al., *Porphyromonas gingivalis and rheumatoid arthritis.* Curr Opin Rheumatol, 2019. **31**(5): p. 517-524.
56. Ebringer, A. and T. Rashid, *Rheumatoid arthritis is caused by a Proteus urinary tract infection.* APMIS, 2014. **122**(5): p. 363-8.
57. Askling, J., et al., *Risk and case characteristics of tuberculosis in rheumatoid arthritis associated with tumor necrosis factor antagonists in Sweden.* Arthritis Rheum, 2005. **52**(7): p. 1986-92.
58. Gilroy, C.B., A. Keat, and D. Taylor-Robinson, *The prevalence of Mycoplasma fermentans in patients with inflammatory arthritides.* Rheumatology (Oxford), 2001. **40**(12): p. 1355-8.
59. Griffiths, D.J., et al., *Detection of human retrovirus 5 in patients with arthritis and systemic lupus erythematosus.* Arthritis Rheum, 1999. **42**(3): p. 448-54.
60. Lidbury, B.A., et al., *Macrophage-derived proinflammatory factors contribute to the development of arthritis and myositis after infection with an arthrogenic alphavirus.* J Infect Dis, 2008. **197**(11): p. 1585-93.

61. Rothe, K., et al., *Latent Cytomegalovirus Infection in Rheumatoid Arthritis and Increased Frequencies of Cytolytic LIR-1+CD8+ T Cells.* Arthritis Rheumatol, 2016. **68**(2): p. 337-46.
62. Chattopadhyay, H., et al., *Demonstration of anti-rubella antibody-secreting cells in rheumatoid arthritis patients.* Scand J Immunol, 1979. **10**(1): p. 47-54.
63. Takahashi, Y., et al., *Human parvovirus B19 as a causative agent for rheumatoid arthritis.* Proc Natl Acad Sci U S A, 1998. **95**(14): p. 8227-32.
64. Balandraud, N. and J. Roudier, *Epstein-Barr virus and rheumatoid arthritis.* Joint Bone Spine, 2017.
65. Humby, F., et al., *Synovial cellular and molecular signatures stratify clinical response to csDMARD therapy and predict radiographic progression in early rheumatoid arthritis patients.* Ann Rheum Dis, 2019. **78**(6): p. 761-772.
66. Lliso-Ribera, G., et al., *Synovial tissue signatures enhance clinical classification and prognostic/treatment response algorithms in early inflammatory arthritis and predict requirement for subsequent biological therapy: results from the pathobiology of early arthritis cohort (PEAC).* Ann Rheum Dis, 2019. **78**(12): p. 1642-1652.
67. Lewis, M.J., et al., *Molecular Portraits of Early Rheumatoid Arthritis Identify Clinical and Treatment Response Phenotypes.* Cell Rep, 2019. **28**(9): p. 2455-2470 e5.
68. Weyand, C.M. and J.J. Goronzy, *Ectopic germinal center formation in rheumatoid synovitis.* Ann N Y Acad Sci, 2003. **987**: p. 140-9.
69. Pitzalis, C., et al., *Ectopic lymphoid-like structures in infection, cancer and autoimmunity.* Nat Rev Immunol, 2014. **14**(7): p. 447-62.
70. Magalhaes, R., et al., *Morphological and molecular pathology of the B cell response in synovitis of rheumatoid arthritis.* Virchows Arch, 2002. **441**(5): p. 415-27.
71. Manzo, A., et al., *Systematic microanatomical analysis of CXCL13 and CCL21 in situ production and progressive lymphoid organization in rheumatoid synovitis.* Eur J Immunol, 2005. **35**(5): p. 1347-59.
72. Barone, F., et al., *Association of CXCL13 and CCL21 expression with the progressive organization of lymphoid-like structures in Sjogren's syndrome.* Arthritis Rheum, 2005. **52**(6): p. 1773-84.
73. Hjelmstrom, P., *Lymphoid neogenesis: de novo formation of lymphoid tissue in chronic inflammation through expression of homing chemokines.* J Leukoc Biol, 2001. **69**(3): p. 331-9.
74. Stebegg, M., et al., *Regulation of the Germinal Center Response.* Front Immunol, 2018. **9**: p. 2469.
75. Bombardieri, M., M. Lewis, and C. Pitzalis, *Ectopic lymphoid neogenesis in rheumatic autoimmune diseases.* Nat Rev Rheumatol, 2017. **13**(3): p. 141-154.
76. Barone, F., et al., *Stromal Fibroblasts in Tertiary Lymphoid Structures: A Novel Target in Chronic Inflammation.* Front Immunol, 2016. **7**: p. 477.

77. Croft, A.P., et al., *Distinct fibroblast subsets drive inflammation and damage in arthritis.* Nature, 2019. **570**(7760): p. 246-251.
78. Muramatsu, M., et al., *Specific expression of activation-induced cytidine deaminase (AID), a novel member of the RNA-editing deaminase family in germinal center B cells.* J Biol Chem, 1999. **274**(26): p. 18470-6.
79. Roco, J.A., et al., *Class-Switch Recombination Occurs Infrequently in Germinal Centers.* Immunity, 2019. **51**(2): p. 337-350 e7.
80. Humby, F., et al., *Ectopic lymphoid structures support ongoing production of class-switched autoantibodies in rheumatoid synovium.* PLoS Med, 2009. **6**(1): p. e1.
81. Edwards, J.C. and G. Cambridge, *B-cell targeting in rheumatoid arthritis and other autoimmune diseases.* Nat Rev Immunol, 2006. **6**(5): p. 394-403.
82. Derksen, V., T.W.J. Huizinga, and D. van der Woude, *The role of autoantibodies in the pathophysiology of rheumatoid arthritis.* Semin Immunopathol, 2017. **39**(4): p. 437-446.
83. Song, Y.W. and E.H. Kang, *Autoantibodies in rheumatoid arthritis: rheumatoid factors and anticitrullinated protein antibodies.* QJM, 2010. **103**(3): p. 139-46.
84. Waaler, E., *On the occurrence of a factor in human serum activating the specific agglutintion of sheep blood corpuscles. 1939.* APMIS, 2007. **115**(5): p. 422-38; discussion 439.
85. Rose, H.M., C. Ragan, and et al., *Differential agglutination of normal and sensitized sheep erythrocytes by sera of patients with rheumatoid arthritis.* Proc Soc Exp Biol Med, 1948. **68**(1): p. 1-6.
86. Dorner, T., et al., *Rheumatoid factor revisited.* Curr Opin Rheumatol, 2004. **16**(3): p. 246-53.
87. Dwivedi, N. and M. Radic, *Citrullination of autoantigens implicates NETosis in the induction of autoimmunity.* Ann Rheum Dis, 2014. **73**(3): p. 483-91.
88. Matsumoto, I., et al., *How antibodies to a ubiquitous cytoplasmic enzyme may provoke joint-specific autoimmune disease.* Nat Immunol, 2002. **3**(4): p. 360-5.
89. Carson, D.A., et al., *Rheumatoid factor and immune networks.* Annu Rev Immunol, 1987. **5**: p. 109-26.
90. Pope, R.M., D.C. Teller, and M. Mannik, *Intermediate complexes formed by self-association of IgG-rheumatoid factors.* Ann N Y Acad Sci, 1975. **256**: p. 82-7.
91. Sebbag, M., et al., *The antiperinuclear factor and the so-called antikeratin antibodies are the same rheumatoid arthritis-specific autoantibodies.* J Clin Invest, 1995. **95**(6): p. 2672-9.
92. Girbal-Neuhauser, E., et al., *The epitopes targeted by the rheumatoid arthritis-associated antifilaggrin autoantibodies are posttranslationally generated on various sites of (pro)filaggrin by deimination of arginine residues.* J Immunol, 1999. **162**(1): p. 585-94.

93. Schellekens, G.A., et al., *Citrulline is an essential constituent of antigenic determinants recognized by rheumatoid arthritis-specific autoantibodies.* J Clin Invest, 1998. **101**(1): p. 273-81.
94. Hill, J.A., et al., *Cutting edge: the conversion of arginine to citrulline allows for a high-affinity peptide interaction with the rheumatoid arthritis-associated HLA-DRB1*0401 MHC class II molecule.* J Immunol, 2003. **171**(2): p. 538-41.
95. Scally, S.W., et al., *A molecular basis for the association of the HLA-DRB1 locus, citrullination, and rheumatoid arthritis.* J Exp Med, 2013. **210**(12): p. 2569-82.
96. de Almeida, D.E., S. Ling, and J. Holoshitz, *New insights into the functional role of the rheumatoid arthritis shared epitope.* FEBS Lett, 2011. **585**(23): p. 3619-26.
97. Ling, S., et al., *Citrullinated calreticulin potentiates rheumatoid arthritis shared epitope signaling.* Arthritis Rheum, 2013. **65**(3): p. 618-26.
98. Corsiero, E., et al., *Characterization of a Synovial B Cell-Derived Recombinant Monoclonal Antibody Targeting Stromal Calreticulin in the Rheumatoid Joints.* J Immunol, 2018. **201**(5): p. 1373-1381.
99. Lundberg, K., et al., *Antibodies to citrullinated alpha-enolase peptide 1 are specific for rheumatoid arthritis and cross-react with bacterial enolase.* Arthritis Rheum, 2008. **58**(10): p. 3009-19.
100. Konig, M.F., et al., *Aggregatibacter actinomycetemcomitans-induced hypercitrullination links periodontal infection to autoimmunity in rheumatoid arthritis.* Sci Transl Med, 2016. **8**(369): p. 369ra176.
101. Cambridge, G., et al., *Serologic changes following B lymphocyte depletion therapy for rheumatoid arthritis.* Arthritis Rheum, 2003. **48**(8): p. 2146-54.
102. Clavel, C., et al., *Induction of macrophage secretion of tumor necrosis factor alpha through Fcgamma receptor IIa engagement by rheumatoid arthritis-specific autoantibodies to citrullinated proteins complexed with fibrinogen.* Arthritis Rheum, 2008. **58**(3): p. 678-88.
103. Laurent, L., et al., *Fcgamma receptor profile of monocytes and macrophages from rheumatoid arthritis patients and their response to immune complexes formed with autoantibodies to citrullinated proteins.* Ann Rheum Dis, 2011. **70**(6): p. 1052-9.
104. Sokolove, J., et al., *Rheumatoid factor as a potentiator of anti-citrullinated protein antibody-mediated inflammation in rheumatoid arthritis.* Arthritis Rheumatol, 2014. **66**(4): p. 813-21.
105. Lu, M.C., et al., *Anti-citrullinated protein antibodies bind surface-expressed citrullinated Grp78 on monocyte/macrophages and stimulate tumor necrosis factor alpha production.* Arthritis Rheum, 2010. **62**(5): p. 1213-23.
106. Bugatti, S., et al., *Anti-citrullinated protein antibodies and high levels of rheumatoid factor are associated with systemic bone loss in patients with early untreated rheumatoid arthritis.* Arthritis Res Ther, 2016. **18**(1): p. 226.

107. Trouw, L.A., et al., *Anti-cyclic citrullinated peptide antibodies from rheumatoid arthritis patients activate complement via both the classical and alternative pathways.* Arthritis Rheum, 2009. **60**(7): p. 1923-31.
108. Winchester, R.J., V. Agnello, and H.G. Kunkel, *Gamma globulin complexes in synovial fluids of patients with rheumatoid arthritis. Partial characterization and relationship to lowered complement levels.* Clin Exp Immunol, 1970. **6**(5): p. 689-706.
109. Pratesi, F., et al., *Antibodies from patients with rheumatoid arthritis target citrullinated histone 4 contained in neutrophils extracellular traps.* Ann Rheum Dis, 2014. **73**(7): p. 1414-22.
110. Corsiero, E., et al., *Single cell cloning and recombinant monoclonal antibodies generation from RA synovial B cells reveal frequent targeting of citrullinated histones of NETs.* Ann Rheum Dis, 2015.
111. Teitsson, I., et al., *Prospective study of early rheumatoid arthritis. I. Prognostic value of IgA rheumatoid factor.* Ann Rheum Dis, 1984. **43**(5): p. 673-8.
112. Aleyd, E., et al., *IgA Complexes in Plasma and Synovial Fluid of Patients with Rheumatoid Arthritis Induce Neutrophil Extracellular Traps via FcalphaRI.* J Immunol, 2016. **197**(12): p. 4552-4559.
113. Hecht, C., et al., *Additive effect of anti-citrullinated protein antibodies and rheumatoid factor on bone erosions in patients with RA.* Ann Rheum Dis, 2015. **74**(12): p. 2151-6.
114. van Steenbergen, H.W., et al., *The effects of rheumatoid factor and anticitrullinated peptide antibodies on bone erosions in rheumatoid arthritis.* Ann Rheum Dis, 2015. **74**(1): p. e3.
115. Wigerblad, G., et al., *Autoantibodies to citrullinated proteins induce joint pain independent of inflammation via a chemokine-dependent mechanism.* Ann Rheum Dis, 2016. **75**(4): p. 730-8.
116. Lindhout, E., et al., *Fibroblast-like synoviocytes from rheumatoid arthritis patients have intrinsic properties of follicular dendritic cells.* J Immunol, 1999. **162**(10): p. 5949-56.
117. Bartok, B. and G.S. Firestein, *Fibroblast-like synoviocytes: key effector cells in rheumatoid arthritis.* Immunological Reviews, 2010. **233**: p. 233-255.
118. Ospelt, C., et al., *Overexpression of toll-like receptors 3 and 4 in synovial tissue from patients with early rheumatoid arthritis: toll-like receptor expression in early and longstanding arthritis.* Arthritis Rheum, 2008. **58**(12): p. 3684-92.
119. Uehara, A. and H. Takada, *Functional TLRs and NODs in human gingival fibroblasts.* J Dent Res, 2007. **86**(3): p. 249-54.
120. Umetsu, D.T., et al., *Antigen presentation by human dermal fibroblasts: activation of resting T lymphocytes.* J Immunol, 1986. **136**(2): p. 440-5.
121. Lee, D.M., et al., *Cadherin-11 in synovial lining formation and pathology in arthritis.* Science, 2007. **315**(5814): p. 1006-10.
122. Firestein, G.S. and I.B. McInnes, *Immunopathogenesis of Rheumatoid Arthritis.* Immunity, 2017. **46**(2): p. 183-196.

123. Swanson-Mungerson, M., R. Bultema, and R. Longnecker, *Epstein-Barr virus LMP2A enhances B-cell responses in vivo and in vitro.* J Virol, 2006. **80**(14): p. 6764-70.
124. Neumann, E., et al., *Cell culture and passaging alters gene expression pattern and proliferation rate in rheumatoid arthritis synovial fibroblasts.* Arthritis Res Ther, 2010. **12**(3): p. R83.
125. Ospelt, C., *Synovial fibroblasts in 2017.* RMD Open, 2017. **3**(2): p. e000471.
126. Bombardieri, M., et al., *A BAFF/APRIL-dependent TLR3-stimulated pathway enhances the capacity of rheumatoid synovial fibroblasts to induce AID expression and Ig class-switching in B cells.* Ann Rheum Dis, 2011. **70**(10): p. 1857-65.
127. Brentano, F., et al., *RNA released from necrotic synovial fluid cells activates rheumatoid arthritis synovial fibroblasts via Toll-like receptor 3.* Arthritis Rheum, 2005. **52**(9): p. 2656-65.
128. Hatterer, E., et al., *A specific anti-citrullinated protein antibody profile identifies a group of rheumatoid arthritis patients with a toll-like receptor 4-mediated disease.* Arthritis Res Ther, 2016. **18**(1): p. 224.
129. Kyburz, D., et al., *Bacterial peptidoglycans but not CpG oligodeoxynucleotides activate synovial fibroblasts by toll-like receptor signaling.* Arthritis Rheum, 2003. **48**(3): p. 642-50.
130. Carmona-Rivera, C., et al., *Synovial fibroblast-neutrophil interactions promote pathogenic adaptive immunity in rheumatoid arthritis.* Sci Immunol, 2017. **2**(10).
131. Carrion, M., et al., *RNA sensors in human osteoarthritis and rheumatoid arthritis synovial fibroblasts: immune regulation by vasoactive intestinal peptide.* Arthritis Rheum, 2011. **63**(6): p. 1626-36.
132. Ospelt, C., et al., *Expression, regulation, and signaling of the pattern-recognition receptor nucleotide-binding oligomerization domain 2 in rheumatoid arthritis synovial fibroblasts.* Arthritis Rheum, 2009. **60**(2): p. 355-63.
133. Yokota, K., et al., *The pattern-recognition receptor nucleotide-binding oligomerization domain--containing protein 1 promotes production of inflammatory mediators in rheumatoid arthritis synovial fibroblasts.* Arthritis Rheum, 2012. **64**(5): p. 1329-37.
134. Girardin, S.E., et al., *Peptidoglycan molecular requirements allowing detection by Nod1 and Nod2.* J Biol Chem, 2003. **278**(43): p. 41702-8.
135. Haniffa, M.A., et al., *Adult human fibroblasts are potent immunoregulatory cells and functionally equivalent to mesenchymal stem cells.* J Immunol, 2007. **179**(3): p. 1595-604.
136. Liddiard, K. and P.R. Taylor, *Understanding local macrophage phenotypes in disease: shape-shifting macrophages.* Nat Med, 2015. **21**(2): p. 119-20.
137. Elshabrawy, H.A., et al., *The pathogenic role of angiogenesis in rheumatoid arthritis.* Angiogenesis, 2015. **18**(4): p. 433-48.

138. Wehmeyer, C., et al., *The role of stromal cells in inflammatory bone loss.* Clin Exp Immunol, 2017. **189**(1): p. 1-11.
139. Pitzalis, C., S. Kelly, and F. Humby, *New learnings on the pathophysiology of RA from synovial biopsies.* Curr Opin Rheumatol, 2013. **25**(3): p. 334-44.
140. Takayanagi, H., et al., *Involvement of receptor activator of nuclear factor kappaB ligand/osteoclast differentiation factor in osteoclastogenesis from synoviocytes in rheumatoid arthritis.* Arthritis Rheum, 2000. **43**(2): p. 259-69.
141. Korb-Pap, A., et al., *Early structural changes in cartilage and bone are required for the attachment and invasion of inflamed synovial tissue during destructive inflammatory arthritis.* Ann Rheum Dis, 2012. **71**(6): p. 1004-11.
142. Vartio, T., et al., *Fibronectin in synovial fluid and tissue in rheumatoid arthritis.* Eur J Clin Invest, 1981. **11**(3): p. 207-12.
143. Rengel, Y., C. Ospelt, and S. Gay, *Proteinases in the joint: clinical relevance of proteinases in joint destruction.* Arthritis Res Ther, 2007. **9**(5): p. 221.
144. Muller-Ladner, U., et al., *Synovial fibroblasts of patients with rheumatoid arthritis attach to and invade normal human cartilage when engrafted into SCID mice.* Am J Pathol, 1996. **149**(5): p. 1607-15.
145. Angiolilli, C., et al., *The acetyl code in rheumatoid arthritis and other rheumatic diseases.* Epigenomics, 2017. **9**(4): p. 447-461.
146. Churov, A.V., E.K. Oleinik, and M. Knip, *MicroRNAs in rheumatoid arthritis: altered expression and diagnostic potential.* Autoimmun Rev, 2015. **14**(11): p. 1029-37.
147. Frank-Bertoncelj, M., et al., *Epigenetically-driven anatomical diversity of synovial fibroblasts guides joint-specific fibroblast functions.* Nat Commun, 2017. **8**: p. 14852.
148. Croft, A.P., et al., *Rheumatoid synovial fibroblasts differentiate into distinct subsets in the presence of cytokines and cartilage.* Arthritis Res Ther, 2016. **18**(1): p. 270.
149. Mizoguchi, F., et al., *Functionally distinct disease-associated fibroblast subsets in rheumatoid arthritis.* Nat Commun, 2018. **9**(1): p. 789.
150. Sidney, L.E., et al., *Concise review: evidence for CD34 as a common marker for diverse progenitors.* Stem Cells, 2014. **32**(6): p. 1380-9.
151. Rege, T.A. and J.S. Hagood, *Thy-1 as a regulator of cell-cell and cell-matrix interactions in axon regeneration, apoptosis, adhesion, migration, cancer, and fibrosis.* FASEB J, 2006. **20**(8): p. 1045-54.
152. Choi, I.Y., et al., *Stromal cell markers are differentially expressed in the synovial tissue of patients with early arthritis.* PLoS One, 2017. **12**(8): p. e0182751.
153. Falconer, J., et al., *Synovial cell metabolism and chronic inflammation in rheumatoid arthritis.* Arthritis Rheumatol, 2018.
154. Epstein, M.A., B.G. Achong, and Y.M. Barr, *Virus Particles in Cultured Lymphoblasts from Burkitt's Lymphoma.* Lancet, 1964. **1**(7335): p. 702-3.

155. Young, L.S., L.F. Yap, and P.G. Murray, *Epstein-Barr virus: more than 50 years old and still providing surprises.* Nat Rev Cancer, 2016. **16**(12): p. 789-802.
156. Sample, J., et al., *Epstein-Barr virus types 1 and 2 differ in their EBNA-3A, EBNA-3B, and EBNA-3C genes.* J Virol, 1990. **64**(9): p. 4084-92.
157. Cohen, J.I., et al., *Epstein-Barr virus nuclear protein 2 is a key determinant of lymphocyte transformation.* Proc Natl Acad Sci U S A, 1989. **86**(23): p. 9558-62.
158. Rickinson, A.B., L.S. Young, and M. Rowe, *Influence of the Epstein-Barr virus nuclear antigen EBNA 2 on the growth phenotype of virus-transformed B cells.* J Virol, 1987. **61**(5): p. 1310-7.
159. Zanella, L., et al., *A reliable Epstein-Barr Virus classification based on phylogenomic and population analyses.* Sci Rep, 2019. **9**(1): p. 9829.
160. Hjalgrim, H., J. Friborg, and M. Melbye, *The epidemiology of EBV and its association with malignant disease*, in *Human Herpesviruses: Biology, Therapy, and Immunoprophylaxis*, A. Arvin, et al., Editors. 2007: Cambridge.
161. Kutok, J.L. and F. Wang, *Spectrum of Epstein-Barr virus-associated diseases.* Annu Rev Pathol, 2006. **1**: p. 375-404.
162. Kuppers, R., *B cells under influence: Transformation of B cells by Epstein-Barr virus.* Nature Reviews Immunology, 2003. **3**(10): p. 801-812.
163. Thorley-Lawson, D.A., *Epstein-Barr virus: exploiting the immune system.* Nature Reviews Immunology, 2001. **1**(1): p. 75-82.
164. Borozan, I., et al., *Analysis of Epstein-Barr Virus Genomes and Expression Profiles in Gastric Adenocarcinoma.* J Virol, 2018. **92**(2).
165. Odumade, O.A., K.A. Hogquist, and H.H. Balfour, Jr., *Progress and problems in understanding and managing primary Epstein-Barr virus infections.* Clin Microbiol Rev, 2011. **24**(1): p. 193-209.
166. Tao, Q., et al., *Epstein-Barr virus (EBV) and its associated human cancers--genetics, epigenetics, pathobiology and novel therapeutics.* Front Biosci, 2006. **11**: p. 2672-713.
167. Ogembo, J.G., et al., *Human complement receptor type 1/CD35 is an Epstein-Barr Virus receptor.* Cell Rep, 2013. **3**(2): p. 371-85.
168. Mohl, B.S., J. Chen, and R. Longnecker, *Gammaherpesvirus entry and fusion: A tale how two human pathogenic viruses enter their host cells.* Adv Virus Res, 2019. **104**: p. 313-343.
169. Nemerow, G.R. and N.R. Cooper, *Early events in the infection of human B lymphocytes by Epstein-Barr virus: the internalization process.* Virology, 1984. **132**(1): p. 186-98.
170. Trier, N., et al., *Human MHC-II with Shared Epitope Motifs Are Optimal Epstein-Barr Virus Glycoprotein 42 Ligands-Relation to Rheumatoid Arthritis.* Int J Mol Sci, 2018. **19**(1).
171. Carel, J.C., et al., *Structural requirements for C3d,g/Epstein-Barr virus receptor (CR2/CD21) ligand binding, internalization, and viral infection.* J Biol Chem, 1990. **265**(21): p. 12293-9.

172. Walling, D.M., et al., *Persistent productive Epstein-Barr virus replication in normal epithelial cells in vivo.* J Infect Dis, 2001. **184**(12): p. 1499-507.
173. Sixbey, J.W., et al., *Epstein-Barr virus replication in oropharyngeal epithelial cells.* N Engl J Med, 1984. **310**(19): p. 1225-30.
174. Pegtel, D.M., J. Middeldorp, and D.A. Thorley-Lawson, *Epstein-Barr virus infection in ex vivo tonsil epithelial cell cultures of asymptomatic carriers.* J Virol, 2004. **78**(22): p. 12613-24.
175. Perry, M.E., M.M. Jones, and Y. Mustafa, *Structure of the crypt epithelium in human palatine tonsils.* Acta Otolaryngol Suppl, 1988. **454**: p. 53-9.
176. Klein, G., et al., *EBV-determined nuclear antigen (EBNA)-positive cells in the peripheral blood of infectious mononucleosis patients.* Int J Cancer, 1976. **17**(1): p. 21-6.
177. Kurth, J., et al., *Epstein-Barr virus-infected B cells expanding in germinal centers of infectious mononucleosis patients do not participate in the germinal center reaction.* Proceedings of the National Academy of Sciences of the United States of America, 2003. **100**(8): p. 4730-4735.
178. Delecluse, H.J., et al., *Episomal and integrated copies of Epstein-Barr virus coexist in Burkitt lymphoma cell lines.* J Virol, 1993. **67**(3): p. 1292-9.
179. Tsurumi, T., M. Fujita, and A. Kudoh, *Latent and lytic Epstein-Barr virus replication strategies.* Rev Med Virol, 2005. **15**(1): p. 3-15.
180. Gordon, L.I. and R. Longnecker, *Off-targeting oft-targeted CD20 in cHL.* Blood, 2012. **119**(18): p. 4095-6.
181. Lestou, V.S., et al., *Non-random integration of Epstein-Barr virus in lymphoblastoid cell lines.* Genes Chromosomes Cancer, 1993. **8**(1): p. 38-48.
182. Reisinger, J., et al., *Visualization of episomal and integrated Epstein-Barr virus DNA by fiber fluorescence in situ hybridization.* Int J Cancer, 2006. **118**(7): p. 1603-8.
183. Thorley-Lawson, D.A. and G.J. Babcock, *A model for persistent infection with Epstein-Barr virus: the stealth virus of human B cells.* Life Sci, 1999. **65**(14): p. 1433-53.
184. Kurth, J., et al., *EBV-infected B cells in infectious mononucleosis: viral strategies for spreading in the B cell compartment and establishing latency.* Immunity, 2000. **13**(4): p. 485-95.
185. Babcock, G.J., et al., *Epstein-barr virus-infected resting memory B cells, not proliferating lymphoblasts, accumulate in the peripheral blood of immunosuppressed patients.* J Exp Med, 1999. **190**(4): p. 567-76.
186. Yates, J.L., N. Warren, and B. Sugden, *Stable replication of plasmids derived from Epstein-Barr virus in various mammalian cells.* Nature, 1985. **313**(6005): p. 812-5.
187. Baer, R., et al., *DNA sequence and expression of the B95-8 Epstein-Barr virus genome.* Nature, 1984. **310**(5974): p. 207-11.
188. Moss, W.N., *Analyses of non-coding RNAs generated from the Epstein-Barr virus W repeat region.* Proceedings Iwbbio 2014:

International Work-Conference on Bioinformatics and Biomedical Engineering, Vols 1 and 2, 2014: p. 238-252.
189. Bodescot, M., et al., *Spliced RNA from the IR1-U2 region of Epstein-Barr virus: presence of an open reading frame for a repetitive polypeptide.* EMBO J, 1984. **3**(8): p. 1913-7.
190. Nitsche, F., A. Bell, and A. Rickinson, *Epstein-Barr virus leader protein enhances EBNA-2-mediated transactivation of latent membrane protein 1 expression: a role for the W1W2 repeat domain.* J Virol, 1997. **71**(9): p. 6619-28.
191. Luftig, M., et al., *Epstein-Barr virus latent membrane protein 1 activation of NF-kappaB through IRAK1 and TRAF6.* Proc Natl Acad Sci U S A, 2003. **100**(26): p. 15595-600.
192. Longnecker, R., *Epstein-Barr virus latency: LMP2, a regulator or means for Epstein-Barr virus persistence?* Adv Cancer Res, 2000. **79**: p. 175-200.
193. Nanbo, A., et al., *Epstein-Barr virus RNA confers resistance to interferon-alpha-induced apoptosis in Burkitt's lymphoma.* EMBO J, 2002. **21**(5): p. 954-65.
194. Komano, J., et al., *Oncogenic role of Epstein-Barr virus-encoded RNAs in Burkitt's lymphoma cell line Akata.* J Virol, 1999. **73**(12): p. 9827-31.
195. Yamamoto, N., et al., *Malignant transformation of B lymphoma cell line BJAB by Epstein-Barr virus-encoded small RNAs.* FEBS Lett, 2000. **484**(2): p. 153-8.
196. Fok, V., et al., *Multiple domains of EBER 1, an Epstein-Barr virus noncoding RNA, recruit human ribosomal protein L22.* RNA, 2006. **12**(5): p. 872-82.
197. Kitagawa, N., et al., *Epstein-Barr virus-encoded poly(A)(-) RNA supports Burkitt's lymphoma growth through interleukin-10 induction.* EMBO J, 2000. **19**(24): p. 6742-50.
198. Church, T.M., et al., *Efficient Translation of Epstein-Barr Virus (EBV) DNA Polymerase Contributes to the Enhanced Lytic Replication Phenotype of M81 EBV.* J Virol, 2018. **92**(6).
199. Coghill, A.E. and A. Hildesheim, *Epstein-Barr virus antibodies and the risk of associated malignancies: review of the literature.* Am J Epidemiol, 2014. **180**(7): p. 687-95.
200. Catalina, M.D., et al., *Differential evolution and stability of epitope-specific CD8(+) T cell responses in EBV infection.* J Immunol, 2001. **167**(8): p. 4450-7.
201. Hislop, A.D., et al., *Epitope-specific evolution of human CD8(+) T cell responses from primary to persistent phases of Epstein-Barr virus infection.* J Exp Med, 2002. **195**(7): p. 893-905.
202. Blake, N., et al., *Human CD8+ T cell responses to EBV EBNA1: HLA class I presentation of the (Gly-Ala)-containing protein requires exogenous processing.* Immunity, 1997. **7**(6): p. 791-802.
203. Orange, J.S., *Human natural killer cell deficiencies and susceptibility to infection.* Microbes Infect, 2002. **4**(15): p. 1545-58.
204. Roughan, J.E., C. Torgbor, and D.A. Thorley-Lawson, *Germinal center B cells latently infected with Epstein-Barr virus proliferate*

extensively but do not increase in number. J Virol, 2010. **84**(2): p. 1158-68.
205. Megyola, C., J. Ye, and S. Bhaduri-McIntosh, *Identification of a sub-population of B cells that proliferates after infection with Epstein-Barr virus.* Virol J, 2011. **8**: p. 84.
206. Lucchinetti, C., et al., *Heterogeneity of multiple sclerosis lesions: implications for the pathogenesis of demyelination.* Ann Neurol, 2000. **47**(6): p. 707-17.
207. Lovett-Racke, A.E., et al., *Decreased dependence of myelin basic protein-reactive T cells on CD28-mediated costimulation in multiple sclerosis patients. A marker of activated/memory T cells.* J Clin Invest, 1998. **101**(4): p. 725-30.
208. Markovic-Plese, S., et al., *CD4+CD28- costimulation-independent T cells in multiple sclerosis.* J Clin Invest, 2001. **108**(8): p. 1185-94.
209. Viglietta, V., et al., *Loss of functional suppression by CD4+CD25+ regulatory T cells in patients with multiple sclerosis.* J Exp Med, 2004. **199**(7): p. 971-9.
210. Ascherio, A., K.L. Munger, and K.C. Simon, *Vitamin D and multiple sclerosis.* Lancet Neurol, 2010. **9**(6): p. 599-612.
211. Ascherio, A. and K.L. Munger, *Epstein-barr virus infection and multiple sclerosis: a review.* J Neuroimmune Pharmacol, 2010. **5**(3): p. 271-7.
212. Balfour, H.H., Jr. and P. Verghese, *Primary Epstein-Barr virus infection: impact of age at acquisition, coinfection, and viral load.* J Infect Dis, 2013. **207**(12): p. 1787-9.
213. Hauser, S.L., et al., *B-cell depletion with rituximab in relapsing-remitting multiple sclerosis.* N Engl J Med, 2008. **358**(7): p. 676-88.
214. Ascherio, A., et al., *Epstein-Barr virus antibodies and risk of multiple sclerosis: a prospective study.* JAMA, 2001. **286**(24): p. 3083-8.
215. Ascherio, A., *Epstein-Barr virus in the development of multiple sclerosis.* Expert Rev Neurother, 2008. **8**(3): p. 331-3.
216. Serafini, B., et al., *Epstein-Barr Virus-Specific CD8 T Cells Selectively Infiltrate the Brain in Multiple Sclerosis and Interact Locally with Virus-Infected Cells: Clue for a Virus-Driven Immunopathological Mechanism.* J Virol, 2019. **93**(24).
217. Pender, M.P., *Infection of autoreactive B lymphocytes with EBV, causing chronic autoimmune diseases.* Trends Immunol, 2003. **24**(11): p. 584-8.
218. Ball, R.J., et al., *Systematic review and meta-analysis of the sero-epidemiological association between Epstein-Barr virus and rheumatoid arthritis.* Arthritis Res Ther, 2015. **17**: p. 274.
219. Tracy, S.I., et al., *Persistence of Epstein-Barr virus in self-reactive memory B cells.* J Virol, 2012. **86**(22): p. 12330-40.
220. Roudier, J., et al., *Susceptibility to rheumatoid arthritis maps to a T-cell epitope shared by the HLA-Dw4 DR beta-1 chain and the Epstein-Barr virus glycoprotein gp110.* Proc Natl Acad Sci U S A, 1989. **86**(13): p. 5104-8.

221. Pratesi, F., et al., *Deiminated Epstein-Barr virus nuclear antigen 1 is a target of anti-citrullinated protein antibodies in rheumatoid arthritis.* Arthritis Rheum, 2006. **54**(3): p. 733-41.
222. Cornillet, M., et al., *In ACPA-positive RA patients, antibodies to EBNA35-58Cit, a citrullinated peptide from the Epstein-Barr nuclear antigen-1, strongly cross-react with the peptide beta60-74Cit which bears the immunodominant epitope of citrullinated fibrin.* Immunol Res, 2015. **61**(1-2): p. 117-25.
223. Harley, J.B., et al., *Transcription factors operate across disease loci, with EBNA2 implicated in autoimmunity.* Nat Genet, 2018. **50**(5): p. 699-707.
224. Ng, K.C., et al., *Anti-RANA antibody: a marker for seronegative and seropositive rheumatoid arthritis.* Lancet, 1980. **1**(8166): p. 447-9.
225. Alspaugh, M.A., et al., *Elevated levels of antibodies to Epstein-Barr virus antigens in sera and synovial fluids of patients with rheumatoid arthritis.* J Clin Invest, 1981. **67**(4): p. 1134-40.
226. Alspaugh, M.A., N. Talal, and E.M. Tan, *Differentiation and characterization of autoantibodies and their antigens in Sjogren's syndrome.* Arthritis Rheum, 1976. **19**(2): p. 216-22.
227. Billings, P.B., et al., *Antibodies to the Epstein-Barr virus nuclear antigen and to rheumatoid arthritis nuclear antigen identify the same polypeptide.* Proc Natl Acad Sci U S A, 1983. **80**(23): p. 7104-8.
228. Merlini, G., et al., *A deiminated viral peptide to detect antibodies in rheumatoid arthritis.* Ann N Y Acad Sci, 2005. **1050**: p. 243-9.
229. Trier, N.H., et al., *Antibodies to a strain-specific citrullinated Epstein-Barr virus peptide diagnoses rheumatoid arthritis.* Sci Rep, 2018. **8**(1): p. 3684.
230. Tosato, G., A.D. Steinberg, and R.M. Blaese, *Defective EBV-specific suppressor T-cell function in rheumatoid arthritis.* N Engl J Med, 1981. **305**(21): p. 1238-43.
231. Toussirot, E., et al., *HLA-DR polymorphism influences T-cell precursor frequencies to Epstein-Barr virus (EBV) gp110: implications for the association of HLA-DR antigens with rheumatoid arthritis.* Tissue Antigens, 1999. **54**(2): p. 146-52.
232. Toussirot, E., et al., *Decreased T cell precursor frequencies to Epstein-Barr virus glycoprotein Gp110 in peripheral blood correlate with disease activity and severity in patients with rheumatoid arthritis.* Ann Rheum Dis, 2000. **59**(7): p. 533-8.
233. Klatt, T., et al., *Expansion of peripheral CD8+ CD28- T cells in response to Epstein-Barr virus in patients with rheumatoid arthritis.* J Rheumatol, 2005. **32**(2): p. 239-51.
234. Balandraud, N., et al., *Epstein-Barr virus load in the peripheral blood of patients with rheumatoid arthritis: accurate quantification using real-time polymerase chain reaction.* Arthritis Rheum, 2003. **48**(5): p. 1223-8.
235. Fujieda, M., et al., *Monitoring of Epstein-Barr virus load and killer T cells in patients with juvenile idiopathic arthritis treated with*

methotrexate or tocilizumab. Mod Rheumatol, 2017. **27**(1): p. 66-71.
236. Miceli-Richard, C., et al., *Effect of methotrexate and anti-TNF on Epstein-Barr virus T-cell response and viral load in patients with rheumatoid arthritis or spondylarthropathies.* Arthritis Res Ther, 2009. **11**(3): p. R77.
237. Bassil, N., et al., *Prospective monitoring of cytomegalovirus, Epstein-Barr virus, BK virus, and JC virus infections on belatacept therapy after a kidney transplant.* Exp Clin Transplant, 2014. **12**(3): p. 212-9.
238. Erre, G.L., et al., *Increased Epstein-Barr Virus DNA Load and Antibodies Against EBNA1 and EA in Sardinian Patients with Rheumatoid Arthritis.* Viral Immunol, 2015. **28**(7): p. 385-90.
239. Croia, C., et al., *Implication of Epstein-Barr virus infection in disease-specific autoreactive B cell activation in ectopic lymphoid structures of Sjogren's syndrome.* Arthritis Rheumatol, 2014. **66**(9): p. 2545-57.
240. Stebulis, J.A., et al., *Fibroblast-like synovial cells derived from synovial fluid.* J Rheumatol, 2005. **32**(2): p. 301-6.
241. Bernasconi, N.L., N. Onai, and A. Lanzavecchia, *A role for Toll-like receptors in acquired immunity: up-regulation of TLR9 by BCR triggering in naive B cells and constitutive expression in memory B cells.* Blood, 2003. **101**(11): p. 4500-4.
242. Aqrawi, L.A., et al., *Autoantigen-Specific Memory B Cells in Primary Sjogren's Syndrome.* Scandinavian Journal of Immunology, 2012. **75**(1): p. 61-68.
243. Torgbor, C., et al., *A multifactorial role for P. falciparum malaria in endemic Burkitt's lymphoma pathogenesis.* PLoS Pathog, 2014. **10**(5): p. e1004170.
244. Corsiero, E., et al., *Generation of Recombinant Monoclonal Antibodies from Single B Cells Isolated from Synovial Tissue of Rheumatoid Arthritis Patients.* Methods Mol Biol, 2018. **1845**: p. 159-187.
245. Butler, A., et al., *Integrating single-cell transcriptomic data across different conditions, technologies, and species.* Nat Biotechnol, 2018. **36**(5): p. 411-420.
246. Stuart, T., et al., *Comprehensive Integration of Single-Cell Data.* Cell, 2019. **177**(7): p. 1888-1902 e21.
247. Osorio, D., et al., *Single-cell RNA sequencing of a European and an African lymphoblastoid cell line.* Sci Data, 2019. **6**(1): p. 112.
248. Szklarczyk, D., et al., *STRING v11: protein-protein association networks with increased coverage, supporting functional discovery in genome-wide experimental datasets.* Nucleic Acids Res, 2019. **47**(D1): p. D607-D613.
249. Boutet, M.A., et al., *Interleukin-36 family dysregulation drives joint inflammation and therapy response in psoriatic arthritis.* Rheumatology (Oxford), 2019.

250. Kaur, M. and L. Esau, *Two-step protocol for preparing adherent cells for high-throughput flow cytometry.* Biotechniques, 2015. **59**(3): p. 119-26.
251. Omi, N., et al., *Efficient and reliable establishment of lymphoblastoid cell lines by Epstein-Barr virus transformation from a limited amount of peripheral blood.* Sci Rep, 2017. **7**: p. 43833.
252. Elliot Sorelle, J.D., Jeffrey Y. Zhou, Stephanie N. Giamberardino, Jeffrey A. Bailey, Simon G. Gregory, Cliburn Chan, Micah A. Luftig, *Single-cell characterization of transcriptomic heterogeneity in lymphoblastoid cell lines.* 2020.
253. Reparon-Schuijt, C.C., et al., *Regulation of synovial B cell survival in rheumatoid arthritis by vascular cell adhesion molecule 1 (CD106) expressed on fibroblast-like synoviocytes.* Arthritis Rheum, 2000. **43**(5): p. 1115-21.
254. Jegerlehner, A., et al., *TLR9 signaling in B cells determines class switch recombination to IgG2a.* J Immunol, 2007. **178**(4): p. 2415-20.
255. Lau, Y.L., et al., *Epstein-Barr-virus-transformed lymphoblastoid cell lines derived from patients with X-linked agammaglobulinaemia and Wiskott-Aldrich syndrome: responses to B cell growth and differentiation factors.* Clin Exp Immunol, 1989. **75**(2): p. 190-5.
256. Hwang, K.K., et al., *Enhanced outgrowth of EBV-transformed chronic lymphocytic leukemia B cells mediated by coculture with macrophage feeder cells.* Blood, 2012. **119**(7): p. e35-44.
257. Yap, H.Y., et al., *Epstein-Barr Virus- (EBV-) Immortalized Lymphoblastoid Cell Lines (LCLs) Express High Level of CD23 but Low CD27 to Support Their Growth.* Adv Virol, 2019. **2019**: p. 6464521.
258. Farina, A., et al., *Epstein-Barr virus infection induces aberrant TLR activation pathway and fibroblast-myofibroblast conversion in scleroderma.* J Invest Dermatol, 2014. **134**(4): p. 954-964.
259. Yoshitomi, H., *Regulation of Immune Responses and Chronic Inflammation by Fibroblast-Like Synoviocytes.* Front Immunol, 2019. **10**: p. 1395.
260. Babcock, G.J., E.M. Miyashita-Lin, and D.A. Thorley-Lawson, *Detection of EBV infection at the single-cell level. Precise quantitation of virus-infected cells in vivo.* Methods Mol Biol, 2001. **174**: p. 103-10.
261. Sanosyan, A., et al., *The impact of targeting repetitive BamHI-W sequences on the sensitivity and precision of EBV DNA quantification.* PLoS One, 2017. **12**(8): p. e0183856.
262. Wiesner, M., et al., *Conditional immortalization of human B cells by CD40 ligation.* PLoS One, 2008. **3**(1): p. e1464.
263. Nanki, T., et al., *Chemokines regulate IL-6 and IL-8 production by fibroblast-like synoviocytes from patients with rheumatoid arthritis.* J Immunol, 2001. **167**(9): p. 5381-5.
264. De Miguel, S., et al., *Mechanisms of CD23 hyperexpression on B cells from patients with rheumatoid arthritis.* J Rheumatol, 2001. **28**(6): p. 1222-8.

265. Mellemkjaer, L., et al., *Rheumatoid arthritis and cancer risk.* Eur J Cancer, 1996. **32A**(10): p. 1753-7.
266. Hong, S., et al., *B Cells Are the Dominant Antigen-Presenting Cells that Activate Naive CD4(+) T Cells upon Immunization with a Virus-Derived Nanoparticle Antigen.* Immunity, 2018. **49**(4): p. 695-708 e4.
267. Kang, S., et al., *Targeting Interleukin-6 Signaling in Clinic.* Immunity, 2019. **50**(4): p. 1007-1023.
268. Horikawa, T., et al., *Induction of c-Met proto-oncogene by Epstein-Barr virus latent membrane protein-1 and the correlation with cervical lymph node metastasis of nasopharyngeal carcinoma.* Am J Pathol, 2001. **159**(1): p. 27-33.
269. Samanta, M., et al., *EB virus-encoded RNAs are recognized by RIG-I and activate signaling to induce type I IFN.* EMBO J, 2006. **25**(18): p. 4207-14.
270. Stuhlmeier, K.M., *Hyaluronan production in synoviocytes as a consequence of viral infections: HAS1 activation by Epstein-Barr virus and synthetic double- and single-stranded viral RNA analogs.* J Biol Chem, 2008. **283**(24): p. 16781-9.
271. Karpus, O.N., et al., *Triggering of the dsRNA sensors TLR3, MDA5, and RIG-I induces CD55 expression in synovial fibroblasts.* PLoS One, 2012. **7**(5): p. e35606.
272. Lai, K.Y., et al., *Maintenance of Epstein-Barr Virus Latent Status by a Novel Mechanism, Latent Membrane Protein 1-Induced Interleukin-32, via the Protein Kinase Cdelta Pathway.* J Virol, 2015. **89**(11): p. 5968-80.
273. Moon, Y.M., et al., *IL-32 and IL-17 interact and have the potential to aggravate osteoclastogenesis in rheumatoid arthritis.* Arthritis Res Ther, 2012. **14**(6): p. R246.