Report on a problem studied at the UK Mathematics-in-Medicine Study Group Cambridge 2014
< https://mmsg.mathmos.net/uk/2014/immunoglobulin-switching/ >

Modelling the regulation of immunoglobulin class switching to IgE and IgG in human B cells to reduce animal use

R. Eljazi, D. Fear, H. Gould, M. Grinfeld, B. Lidbury, V. Manhas, N. Monk, M. Tindall

November 24, 2014

Contents

1 Statement of the problem

The production of high affinity immunoglobulins of specific isotypes, tailored to respond to different immunologic challenges, is central to the success of the vertebrate adaptive immune system and must be tightly regulated to maintain health. For example, the production of IgE is required to clear helminth infections, yet results in allergy when B cells recognising "innocuous" antigens switch to produce IgE in the nose (hayfever), lung (asthma) or skin (eczema). Similarly, the dysregulated production of IgG lies at the heart of many immune conditions such as arthritis and autoimmune disease. Upon antigen encounter in the "germinal centres" of lymph nodes (see Figure 1), B cells can change the isotype and affinity of the immunoglobulin they produce through processes known as class switch recombination (CSR) [1] and somatic hypermutation (SHM) [2].

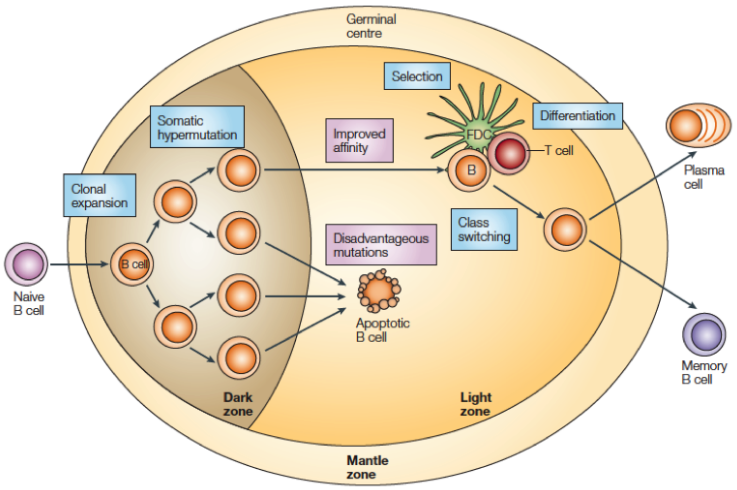

Figure 1: **Cellular process occurring during the germinal centre reactions [3]. Germinal centre reactions begin when naïve B cells encounter antigen in the lymph node (in the presence of T cells or Follicular Dendritic cells) and start to proliferate, forming the germinal centre structure. Rapid B cell proliferation (accompanied by Somatic mutation of the Ig genes) occurs in the dark zone. Following proliferation, B cells migrate to the light zone where they are selected for antigen binding by interaction with T cells and FDCs, rescuing them from apoptosis. Interaction with T helper cells induces class switching and maturation of the B cell into Ig secreting plasma cells or long-lived memory B cells.**

Thus the germinal centre reactions require a carefully choreographed series of events to take place including: gene transcription, DNA breakage and recombination, cell division, regulation of apoptosis and cell differentiation into either antibody producing plasma cells or long-lived memory B cells. To date, little is known about the molecular mechanisms that regulate these activities, target CSR to the IgE gene or how this process may be subverted in asthma and allergy.

Class switch recombination takes place when B cells receive cytokine signals in the presence of T-cell help, initiating the transcription of the "germline" immunoglobulin heavy-chain genes and activation of the recombination pathway. Immunoglobulin class switch recombination to IgE can occur either directly from IgM or via IgG as an intermediate and it is thought that the "choice" of these different pathways has repercussions for the affinity of the resultant antibody and longevity of the B cells produced [4, 5]. It is known that cell division is required for the completion of class switching, and it has been proposed that

division number may regulate the outcome of the process [6]. However, the basis for this division- linked regulation is still poorly understood. Recently, much attention has been given to the mathematical modelling of various aspects of the germinal centre reactions [7, 8, 9] and we feel that a similar investigation of the mechanisms directing class switching would be highly informative.

Our goal for the study group was to model the kinetics of these reactions to understand which might be the rate limiting steps are and why production of IgE occurs less frequently than IgG. More precisely, we wanted to understand what modelling assumptions were necessary to explain the experimental findings shown in Figures 2 and 3.

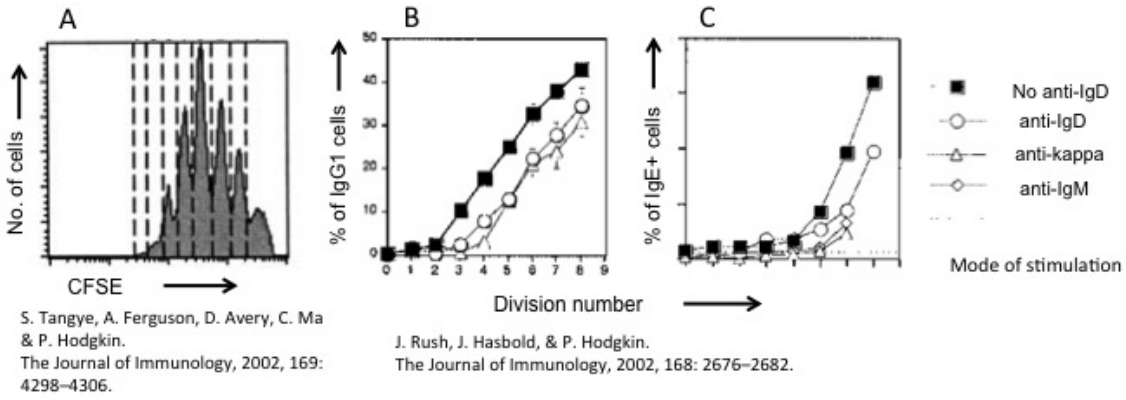

Figure 2: **B cell proliferation dynamics and class switching. B cells have previously been studied by staining cells with CFSE (Carboxyfluorescein succinimidyl ester) (reviewed in [6]), yielding reproducible proportions of cells that have undergone a fixed number of divisions (A) and suggesting that a defined number of divisions is required to switch to different isotypes (B&C).**

In addition, we wished to interrogate genome wide expression data that we had previously collected in a more "unbiased" manner with a view to identifying factors that might regulate and particularly co-ordinate the processes described above. To date much research in this area has been carried out in mice, utilising *in vitro* stimulation of primary murine B cells (isolated post-mortem from the spleen and bone marrow) and mouse models of asthma and allergic disease. These studies use large numbers of animals and are associated with animal welfare concerns; asthma disease models require repeated dosing of inhaled allergen over the course of several weeks and long term kinetic investigations of circulating and lung lavage cell numbers

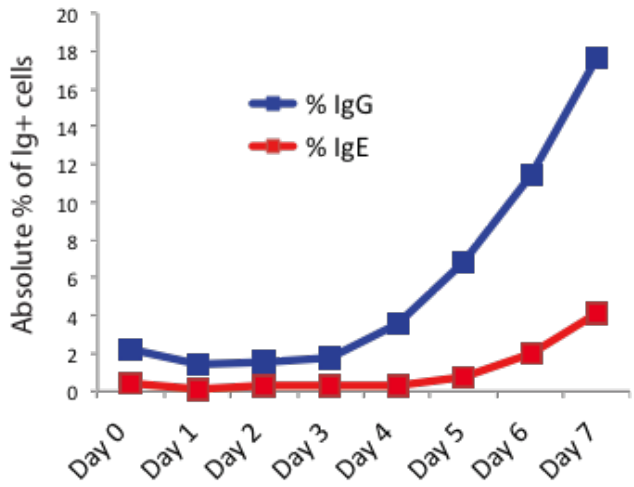

Figure 3: **Dynamics of the appearance of class switched B cells in in-vitro cultures. The appearance of class switched (IgG+ and IgE+) B cells was followed by flow cytometry over the time course of the in-vitro culture following IL-4 and anti-CD40 stimulation.**

over the course of months in mice with impaired life expectancy. However, we (and others) have shown that class switching appears to be regulated differently in mice vs. man and that as such mouse models of disease have not been predictive, particularly regarding asthma (H. Gould Trends in Immunology, in press). We have previously established a human *ex vivo* culture system that we believe models these germinal centre reactions more closely than the murine system and has more relevance to investigations that could advance human healthcare. We hope that by the application of the methods described above we can demonstrate the utility of this model and identify targets that we can take forward into mechanistic studies.

2 Transcriptional profiling of IgM to IgG/IgE switching in human B-lymphocytes post interleukin-4 stimulation via recursive partitioning

2.1 Introduction

To provide biological insights into B-cell activity post interleukin-4 (IL-4) stimulation, gene transcript expression data collected from cultures were interogated by two recursive partitioning methods, single decision trees and the multiple decision tree method, random forests [11, 12].The application of tree algorithms to transcript data allowed for the identification of the statistically most important gene transcripts stimulated

post IL-4 exposure (from approximately 20,000 transcripts analysed, 4,000 were significantly up or down regulated). Of the 4,000 initially significant IL-4-stimulated transcripts, 30 transcripts were identified that spanned cell cycle, cell death (apoptotic) and immune pathways, with subsequent tree analysis revealing the most important transcripts.

2.2 Methods

Initial screening of the gene transcription data was performed by a Bonferroni correction for multiple hypothesis testing. All transcripts had a *p* - value calculated to estmate the significance of transcript fold increase or decrease post IL-4, versus the control non-stimulated cultures (day 0). The Bonferroni correction B was calculated according to $B = \alpha/n$, where $\alpha = 0.05$ (significance level for a single hypothesis), and n is the number of hypotheses tested. For the data set provided, there were $\approx 16,400$ transcripts tested, and 5 time conditions tested versus day 0 (days 0.5, 1.5, 3, 5, 12 post IL-4), resulting in the number of hypotheses estimated at approximately 82,000. From this calculation a p value of 6.1×10^{-7} was estimated. Based on this calculation, only transcripts with $p \leq 10^{-8}$ were considered for further analysis by tree methods.

2.2.1 Single and multiple recursive partitioning investigation of top 30 gene transcripts

Recursive partitioning, known commonly as decision trees, is a machine learning technique for the prediction of class labels from a finite set of data, and was used here as a method to predict gene transcript importance against the time variable (days 0.5 - 12) post IL-4 stimulation. Single decision trees also provide the advantage of calculating single or multiple predictor variable thresholds associated with the response of interest.

The algorithm relies upon calculations of impurity, either via entropy $-\sum_{k=1}^{m} p_i \log p_i$, or the Gini index $1 - \sum_{k=1}^{m} p_i^2$, followed by the application of impurity measures, given by

$$\sum_{j=1}^{k} \left(|E_j|/|E|\right) I(E_j) \tag{1}$$

to derive k possible answers by dividing E into subsets $\{E_1, \ldots, E_k\}$ with minimal mean weighted impurity (1) [11, 12]. Here E is the number of total training items, m is the number of classification classes, p_i is the

fraction of E items that belong to E_i, k is the number of possible answers once impurity I is considered and the weighted average of I is minimised for ultimate best predictor variables ("children nodes").

Thirty gene transcripts with statistical significance for fold change versus day 0 control cultures of $p < 10^{-8}$ (Table 1) were analysed by single decision trees and random forest (1000 individual trees; 10 random predictor variables tested per tree). The weighted importance of each transcript from Table 1 was decided from a mean decrease Gini Index with a weighted score of 12 indicating the most important transcript, to zero indicating no importance of the transcript to IL-4-induced expression over time.

2.3 Results

Bonferroni correction identified 30 gene transcripts with statistical significance of $p < 10^{-8}$ (actual significance level was calculated as 6.1×10^{-7}) (Table 1; the data is from human *ex vivo* B-cell cultures from primary lymphoid tissue (tonsil). Data collected at 0.5, 1.5, 3, 5 and 12 days post IL-4 stimulation. Change in gene expression compared to resting cultures (no IL-4) at day 0), which were thereafter interogated by recursive partitioning analyses. These transcripts reflect the most significant general increase or decrease of specific gene expression in primary human B-cells at any post IL-4 exposure timepoint (days 0.5, 1.5, 3, 5, 12) in comparion to control cultures without IL-4 stimulation. The profile of highly significant transcripts included immune gene expression (e.g. CCL17), transcription machinery (e.g. BATF3), chromatin remodelling (e.g. HIST1H1C), apoptotic and cell cycle factors, as well as a newly characterised and potentially novel factor with with rôles in cytoskeleton function and mitotic divison (MIIF).

Table 1: **All B-lymphocyte gene transcripts with significant up or down expression at $p < 10^{-8}$ at any day post IL-4 stimulation (days 0.5 -12)**

Gene Name (Abbreviation)	Transcript ID (Gene assignment)	Chromosome location	Function/Description
MIIP	2320657 (NM_021933)	1p36.22	Migration & invasion inhibitory protein
TNFSF4	2444283 (NM_003326)	1q25	Tumor Necrosis Factor (ligand) superfamily, member 4
KIF14	2450345 (NM_014875)	1q32.1	Kinesin family member 14
BATF3	2454818 (NM_018664)	1q32.3	Basic leucine zipper transcription factor, ATF-like 3
ARID5A	2494537 (NM_212481)	2q11.2	AT rich interactive domain 5A (MRF1-like)
IKZF2	2597867 (NM_1079526)	2q34	IKAROS family zinc finger 2 (Helios)
BHLHE40	2608725 (?)	3p26	Basic helix-loop-helix family, member e40
IL17RB	2624565 (NM_018725)	3p21.1	Interleukin 17 receptor B
EPHB1	2643592 (NM_004441)	3q21-q23	EPH receptor B1

MARCH1	2792166 (?)	?	Membrane-associated ring finger (C3HC4) 1, E3 ubiquitin
AHRR	2798586 (NM_020731)	5p15.3	Aryl-hydrocarbon receptor repressor
ATXN1	2943434 (NM_000332)	6p23	Ataxin 1
HIST1H1C	2946232 (NM_005319)	6p21.3	Histone cluster 1, H1c
RPS6KA2	2984655 (NM_021135)	6q27	Ribosomal protein S6 kinase, 90kDa, polypeptide 2
TFEC	3069082 (NM_012252)	7q31.2	Transcription factor EC
FGD3	3179551 (NM_1083536)	?	FYVE, RhoGEF and PH domain containing 3
NFIL3	3214451 (NM_005384)	9q22	Nuclear factor, interleukin 3 regulated
TRAF1	3223738 (?)	9q33-q34	TNF receptor-associated factor 1
KIF18A	3367338 (NM_031217)	11p14.1	Kinesin family member 18A
GPR18	3522644 (NM_00529)	13q32	G protein-coupled receptor 18
HOMER2	3636391 (NM_199330)	15q24.3	Homer homolog 2 (Drosophila)
CCL22	3662687 (NM_002990)	16q13	Chemokine (C-C motif) ligand 22
CCL17	3662710 (NM_002987)	16q13	Chemokine (C-C motif) ligand 17
EBI3	3817380 (NM_005755)	19p13.3	Epstein-Barr virus induced 3
TRIP10	3818515 (NM_004240)	19p13.3	Thyroid hormone receptor interactor 10
BCL2L1	3902489 (NM_138578)	20q11.21	BCL2-like 1
B4GALT5	3908963 (NM_004776)	?	UDP-Gal:betaGlcNAc beta 1,4-galactosyltransferase, polypeptide
TOM1	3944084 (NR_024194)	22q13.1	Target of myb1 (chicken)
XBP1	3956589 (NM_005080)	22q12.1	X-box binding protein 1

Subsequent multiple decision tree analysis (random forest) provided an overall importance weighting (mean decrease Gini Index; scale $0 - 12$) of specific transcripts relative to time post IL-4 stimulation (Table 2). Prominent among the most important transcripts were the transcriptional regulator proteins FGD3, NFIL3, IKZF2 and BATF3, with MIIP weighted as the second most important transcript stimulated in human B-cells post IL-4. Immune associated proteins, for example TNFSF4, TRAF1 and CCL17, also featured transcripts of importance (Table 2), but with weighted importance scores < 5 on the Gini Index scale. Single decision trees consistently found MIIP, FGD3 and TNFSF4 as the leading transcripts associated with IL-4 stimulation of human B-cell cultures, with fold change thresholds detected at $((><)6.0$ to $(<)9.0$ for predictions of time (days) post IL-4 (results not shown). MIIP is of particular interest due to its novelty, as well as for its dual rôle for cytoskeleton kinetics and mitotic division regulation.

Table 2: **Summary of top ranked transcripts detected in IL-4 stimulated primary human B-cell cultures. Results reflect the impact of IL-4 stimuation over time (resting time 0 versus days 0.5, 1.5, 3, 5 and 12 post stimulation).**

Array ID Code	Gene or Protein Name	Chromosome location + (Close loci, function)	Importance (Mean Decrease Gini Index)[1]
3179551	FGD3 // FYVE, RhoGEF and PH domain containing 3	?	> 10
2320657	MIIP // migration and invasion inhibitory protein	1p36.22 (proximal to TNFRSF8)	> 8
3214451	NFIL3 // nuclear factor, interleukin 3 regulated	9q22 (Transcription - binds to ATF)	$7.5 - 8$
3367338	KIF18A // kinesin family member 18A	11p14.1 (kinesin superfamily of microtubule-associated molecular motors)	$7.5 - 8$
3908963	B4GALT5 // UDP-Gal:betaGlcNAc beta 1,4-galactosyltransferase	20q13.1-q13.2 (directs protein to Golgi apparatus)	7.5
2597867	IKZF2 // IKAROS family zinc finger 2 (Helios)	2q34 (haematopoietic-specific transcription factor - reg lymphocyte development)	6.0 - 6.5
2454818	BATF // basic leucine zipper transcription factor, ATF-like 3	1q32.3 (transcriptional repressor; IL-2 repressor?)	6.0
Other top 30 transcripts of interest			
2444283	TNFSF4 // tumor necrosis factor (ligand) superfamily, member 4	1q25 (prox - TNFSF18. Polymorphism assoc with SLE)	4.0 - 4.5
3223738	TRAF1 // TNF receptor-associated factor 1	9q33-q34 Prox – C5. Activate NFκB. TRAF1 + TRAF2 = anti-apoptotic	1.5 - 2.0
3662710	CCL17 // chemokine (C-C motif) ligand 17	16q13 (Prox to CCL22 and cytokine-induced apoptosis inhibitor 1. T-cell rôle only)	1.0
3662687	CCL22 // chemokine (C-C motif) ligand 22	16q13 (See CCL17)	< 1.0

With several gene transcripts identified and importance values calculated, subsequent analyses considered the significance of individual transcripts at each time point post IL-4 stimulation.

Table 3 summarises the pattern of significant ($p \leq 10^{-8}$) up or down regulation of key B-cell transcripts at each of the post IL-4 time points where RNA was collected for microarray. While all transcripts were initially brought to attention for significance (after Bonferroni correction) against time in general, closer scrutiny showed that the time of significant increase or decrease of gene expressionwas variable, depending on the gene in question. MIIF showed prominence in both single tree and forest analysis, but was significant only at day 12 post stimulation. The immune gene transcripts TRAF1, CCL17 and CCL22 showed significant increases in expression at all time points post IL-4, with IL-3-regulated nuclear factor (NFIL3) also significantly and positively expressed at each post IL-4 time point. Another prominent transcript detected by tree analysis, FGD3, was significantly downregulated in expression at days 0.5, 1.5 and 3 post stimulation, but not for later time points. These specific transcript differences in significant expression patterns will be of

[1]Highest transcript importance scored at ≥ 10 on the Mean Decrease Gini Index; lowest importance equals 0 on the same scale. RandomForest analysis run on R (v3.0.0); 1000 trees ($ntree = 1000$) with 10 transcripts randomly tested per tree ($mtry = 10$).

value in tandem with predictions from the mathematical modelling of cell growth and cell death phases after exposure to IL-4. Further understanding of transcript activity at individual time points is presented in Figure 4, where fold differences are summarised and individual transcript kinetics emphasised, complementing the findings summarised in Table 3.

Table 3: **Pattern of significant ($p < 10^{-8}$) transcription increase/decrease for 25 genes from human *ex vivo* B-cell cultures stimulated by IL-4. Significance determined by fold increase versus day 0 (pre-IL-4) for 0.5, 1.5, 3, 5 and 12 days post-stimulation. —: Transcript downregulated. Grey shading means significant up or down-regulation at the indicated time point.**

Gene Transcript	Days of Significant Activity Post IL-4 Stimulation				
	0.5	1.5	3	5	12
AHRR					
ARID5A	—	—	—		
ATXN1					
B4GALT5					
BATF3					
BCL2L1					
BHLHE40					
CCL17					
CCL22					
EBI3					
EPHB1					
FGD3	—	—	—		
GPR18	—	—			
HOMER2					
IL17RB					
KIF14					
MARCH1	—		—	—	—
MIIP					
NFIL3					
RPS6KA2					
TFEC					
TNFSF4					
TOM1					
TRAF1					
TRIP10					
XBP1					

2.4 Discussion

1. Literature search to determine whether any of the above identified proteins have been associated with B-cell activation etc post IL-4, as well as in relation to allergy, particularly IgE production and IgE vs IgG switching;

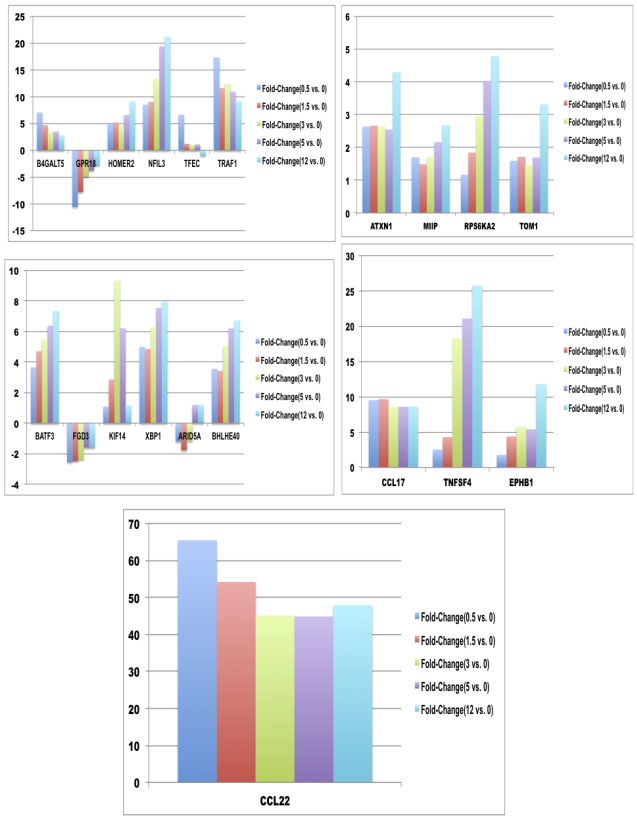

Figure 4: **Gene transcript expression patterns selected by relative importance on recursive partitioning entropy scales and Gini index (see Table 2). Results represent positive or negative fold change post IL-4 versus non-stimulated primary human B-cell control cultures (no IL-4). Only transcripts with statistical significance at $p \leq 10^{-8}$ post Bonferroni correction were considered.**

2. Next steps: specific qT-PCR, ELISA, other protein identification methods, on repeat prospective human B-cell cultures - have the statistical + machine learning predictions been validated in the biological system?

3. Limitations: for tree/forest analysis we need dozens to thousands of cases. Due to expense and other reasons, only limited transcript data was available with rudimentary adjustments made for this study. If new microarrays, RNA seq cannot be conducted, new analysis can be performed after data "bootstrapping", "bagging" or similar sampling methods on the original small data set to enhance predictions through randomly generating new cases.

3 Cell level modelling

3.1 Introduction

We have considered number of different modelling frameworks. The most general of these is a 22-dimensional structured population model, which we will describe first; a systematic analysis of it is left to future work as are simulations using a Gillespie algorithm-based stochastic model. Here we present an analysis of 3-dimensional lumped ODE model and an explicit solution of a birth-death process based on an individual-based model of [10].

3.2 A structured population model

3.2.1 Introduction

Given the experimental results, the most general framework that can be used to model them is a structured population model that takes into account that transitions among different cohorts of cells. Such a model would be couched in a system of ODEs. This system of ODEs can be supplemented by equations for the kinetics of soluble factors such as IL-4 and anti-CD40. Looking to the future, since aggregation of cells is a prominent feature of the *in vitro* set-up, it is desirable to take into account cell motility, as well diffusion of other soluble factors such as immunoglobulin molecules shed by cells producing them.

In this part of the report we mainly discuss the dynamics of cell cohorts, as this has been the thrust of the biologists' questions. We start by writing the most general form of the equation that would be satisfied by the density of a cohort and then specialise to the case of realistic cohorts disregarding any diffusion effects.

3.2.2 The most general form of the equations

We only write such an equation for a cell cohort. Assume that there are N different cell cohorts (distinguished by the number of divisions they have been through and the type of immunoglobulin they produce). Denote the density of the i-th cohort at time t at location x of the domain D in which the cells live (typically, D is a circular domain corresponding to the Petri dish) by $c_i(x,t)$, and the N-vector of all cell densities by $\boldsymbol{c} = (c_1, \ldots c_N)$. Denote the vector of concentrations of all soluble factors by $\boldsymbol{s}(x,t)$; let its dimension be M. Typically this vector would contain as components concentrations of the three immunoglobulins, of IL-4 and of anti-CD40. The the most general equation for the dynamics of c_i has the form

$$\frac{\partial c_i}{\partial t} = \nabla \cdot (D_i(\boldsymbol{s}, \boldsymbol{c}) \nabla c_i) + \sum_{k=1, k\neq i}^{N} f_{k,i}(c_k, \boldsymbol{s}) - \sum_{k=1, k\neq i} f_{i,k}(c_i, \boldsymbol{s}) - m_i(c_i, \boldsymbol{s}).$$

Here $D_i(\boldsymbol{c}, \boldsymbol{s})$ is a hold-all diffusion term that expresses the dependence of the diffusion coefficient on soluble facts and other cohorts; $f_{l,n}$ is the rate of moving from cohort l to n (either by division; see below, or by shift in the type of immunoglobulin produced, and m_i is the mortality rate of cells in cohort i.

Such an equation can be written for each cohort; these N equations are to be supplemented by M equations for soluble species. Working out functional relations is not easy. Once that, and order of magnitude estimates for constants are available, the resulting $N + M$ in general quasilinear reaction-diffusion equations are to be solved on D with Neumann (no-flux) boundary conditions. A rough estimate is that $M \approx 5$ and $N \approx 22$, so the hope of obtaining analytic results is slim.

Trying to specify such a model seems to us nonetheless a useful exercise, as it immediately raises questions of factual ignorance that need to be investigated: e.g. how does the death rate of cells depend on what the cells are doing, how much IL-4 is available, and so on. Similarly, specifying the dynamics of IL-4 requires knowing the dynamics of IL-4 binding and internalisation by the different cohorts of cells.

3.2.3 The simplest possible structured population model of the *in vitro* system

The simplifying assumption we make here is that diffusion mechanism are not being taken into account and that the concentrations of the soluble species are either constant in time or play no rôle.

Furthermore, we will assume that **all interactions are linear**. We will denote densities of cells producing immunoglobulin of type IgX by X. We will take into account cells having undergone up to 8 divisions. This

means that we have the following cohorts:

1. IgM-producing cells: $\{M_0, \ldots, M_8\}$;
2. IgG-producing cells: $\{G_2, \ldots, G_8\}$;
3. IgE-producing cells: $\{E_3, \ldots, E_8\}$.

It is tedious to write down all the equations. For example, the equation for E_4 would be

$$\frac{d}{dt}E_4 = f_{M_4,E_4}M_4 + f_{G_4,E_4}G_4 + f_{E_3,E_4}E_3 - f_{E_4,E_5}E_4 - m_{E_4}E_4.$$

Note that f_{M_4,E_4} is a CSR rate, while f_{E_3,E_4} is a division rate! Similarly, the equation for G_4 would be

$$\frac{d}{dt}G_4 = f_{M_4,G_4}M_4 - f_{G_4,E_4}G_4 + f_{G_3,G_4}G_3 - f_{G_4,G_5}G_4 - m_{G_4}G_4.$$

Let us simplify the system further. Let us assume that the division rate is constant across the cohorts, and that so is the mortality rate. Let us call the common division rate d, and the common mortality rate m, so that the E_4 equation becomes

$$\frac{d}{dt}E_4 = f_{M_4,E_4}M_4 + f_{G_4,E_4}G_4 + dE_3 - dE_4 - mE_4.$$

Now the question asked by the biologists can be formulated as follows: what specifications of f_{M_i,E_i}, f_{M_i,G_i} and f_{G_i,E_i} give plots of **proportions** of IgG- and IgE-producing cells versus time as obtained in experiments?

The consensus among biologists is that assuming that these rates are independent of i does not give the right results, and that each of the rates f must be an increasing function of i; we comment on this below.

Two remarks are in order. First of all the 22-dimensional system of linear ODEs can be explicitly solved. This is because the equations for M_0 decouples and one can sequentially solve then all the M equations; once these are solved, one can solve the G equations, and once that is done, one can solve the E equations.

Second, one can make the assumption of a particular form, of dependence of the fs on i, for example:

$$f_{M_i,E_i} = H(i-3)\alpha_{M,E}(1 - e^{\beta_{M,E}i}).$$

Such an Ansatz reduces the number of unknown constants associated with CSR to only 6.

3.3 A 3-dimensional ODE model of immunoglobulin class switching

A gross simplification of the model detailed in Section 3.2.3 is to consider a 3-dimensional model of the immunoglobin switching process as shown in Figure 5. This allows us to undertake mathematical analysis of the system given the reduced complexity of the governing system of ODEs.

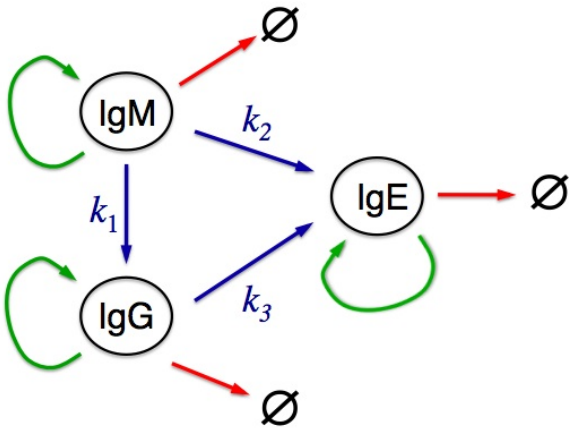

Figure 5: **The three population model**

Assuming cells occupy either one of the three states IgM, IgG and IgE, of which their population densities are denoted M, G and E, respectively, and proliferation and death occur exponentially, we arrive at

$$\frac{\mathrm{d}M}{\mathrm{d}t} = (\alpha_M - k_1 - k_2 - d_M)M = \beta M, \tag{2}$$

$$\frac{\mathrm{d}G}{\mathrm{d}t} = \alpha_G G + k_1 M - k_3 G - d_G G, \tag{3}$$

$$\frac{\mathrm{d}E}{\mathrm{d}t} = \alpha_E E + k_2 M + k_3 G - d_E E \tag{4}$$

where α_X and d_X with $X \in (M, G, E)$ represent the respective rate constants of cell proliferation and death and k_1, k_2 and k_3 are the transistion rate constants for switching between the IgM and IgG, IgM and IgE and IgG and IgE states, respectively. Here $\beta = \alpha_M - k_1 - k_2 - d_M$.

The system is closed with the initial conditions

$$M = M_0, \quad G = 0 \quad \text{and} \quad E = 0, \tag{5}$$

which states that all cells initially occupy the IgM state.

3.3.1 Model analysis

Equations (2)-(4) constitute a system of linear ODEs for which explicit solutions can be obtained. Equation (2) de-couples from the remaining two and can be integrated to give

$$M(t) = M_0 e^{\beta t}. \tag{6}$$

Substituting this result into equation (3) and integrating yields

$$G(t) = \frac{k_1 M_0}{(\beta - \gamma)} \left(e^{\beta t} - e^{\gamma t} \right), \tag{7}$$

where $\gamma = \alpha_G - k_3 - d_G$. Subsequently substituting this result into equation (4) and integrating yields

$$E(t) = \left[\frac{k_1}{(\beta - a)} + \frac{k_1 k_3}{(\beta - \gamma)(\beta - a)} \right] M_0 e^{\beta t} - \frac{k_1 k_3 M_0}{(\beta - \gamma)(\gamma - a)} e^{\gamma t} + \frac{M_0}{(\beta - a)} \left[\frac{k_1 k_3}{(\gamma - a)} - k_2 \right] e^{at}, \tag{8}$$

where $a = \alpha_E - d_E$.

We further note that in the case of equal proliferation rates ($\alpha = \alpha_M = \alpha_G = \alpha_E$) and death rates ($d = d_M = d_G = d_E$) that on addition of equations (2)-(4) we obtain

$$\frac{\mathrm{d}N}{\mathrm{d}t} = (\alpha - d)N \qquad \rightarrow \qquad N(t) = N_0 e^{(\alpha - d)t},$$

where $N = M + G + E$. Thus the total population is dependent upon the rates of proliferation and death at each stage as expected.

3.3.2 Parameterisation

We consider an initial cell population comprising of 100 IgM cells. In order to parameterise the rate constants in the model our group determined the average rate of apoptosis (cell death) for each cell type to be approximately 0.1/day, thus $d_M = d_G = d_E = 0.1$/day. From this we assumed the rate of cell proliferation would be considerably greater in order to maintain a healthy cell population, i.e. we assumed $\alpha_M = \alpha_G = \alpha_E = 0.5$/day. Biologically we are interested in considering how each Ig population varies in time for the case where the rate of transistion from IgM to IgG dominates over those of IgM to IgE and IgG to IgE, which are considered approximately equal, i.e. $k_1 > k_2 \sim k_3$. Here we let $k_1 = 0.3$/day and $k_2 = 0.05$/day$= k_3$/day.

3.3.3 Results

Figure 6 demonstrates the variation in the IgM, IgG and IgE populations over a period of 10 days. Here we see that there is an exponential like increase in IgG and IgE, but IgM remains relatively static over this period. This is explained by the fact that the rates of transition out of the IgM state (cell death and transition to IgG and IgE) are quite rapid in comparison to how fast cells are produced, such that the population remains relatively constant during this period ($\beta \ll 1$). In comparison IgM cells transit quite rapidly to the IgG state. Here the transition to IgE is much less such that the IgG population now has time to increase via proliferation. This increase is subsequently passed onto the IgE population, albeit at a later time than that of the IgG following the transition of cells to that state. This result is made further clear by considering how the transition to IgE from IgM and IgG varies in time as shown in Figure 7.

We next wanted to see what influence the transition rates alone, i.e. in the absence of cell proliferation and cell death, had on the variation of each population in time. As such we set the proliferation rate constants and the death rate constants to be equal ($\alpha_M = \alpha_G = \alpha_E = 0.5/\text{day}=d_M = d_G = d_E$). This led to the results shown in Figures 8 and 9. In Figure 8 we see that the IgM population decays as cell transit to either the IgG or IgE state. There is an initial rapid increase in the IgG population as cells quickly enter that state, before moving onto the IgE one. The IgG cell population starts to decrease around day 6 to 7 as cells leave that state for IgE with no further increases from cell proliferation. Similar to the previous case in which cell proliferation dominated over cell death, the majority of cells entering the IgE state are from IgG as seen in Figure 9; again a result of rapid transit to IgG from IgM and a slower transition directly from IgM.

3.3.4 Conclusions and future work

We have developed a simple linear ODE model of immunogloblin switching between the three states of IgM, IgG and IgE. The model includes a basic exponential description of cell proliferation and death and has been parameterised using currently available data and informed parameter estimates therefrom. For the current given parameter set the model results show how cell transition rates and respectively cell proliferation (when dominating over cell death) affect the transition to IgE from IgM and IgG; when the IgM to IgG transition occurs more rapidly than that of IgM to IgE, cells have time to accummulate in the IgG state, which when combined with a comparatively high proliferation rate leads to a large increase in the IgG and subsequently

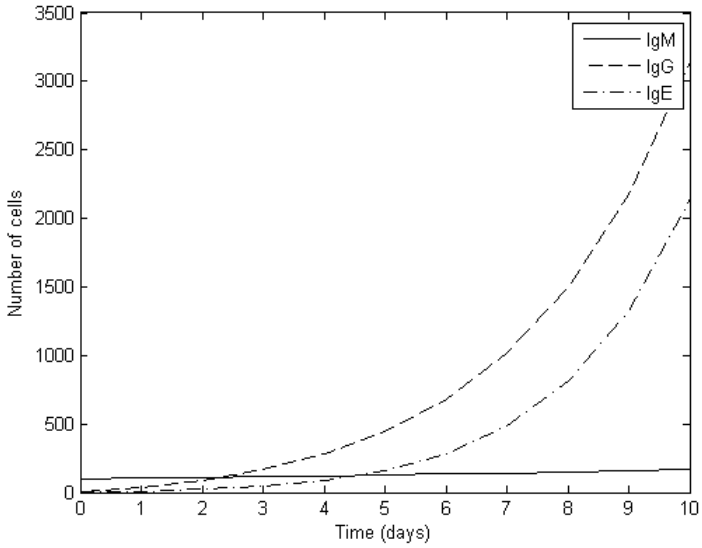

Figure 6: **The variation in the IgM, IgG and IgE populations over a 10 day period as predicted by equations (2)-(4).** Here $\alpha_M = \alpha_G = \alpha_E = 0.5/\text{day}$, $d_M = d_G = d_E = 0.1/\text{day}$, $k_1 = 0.3/\text{day}$ and $k_2 = 0.05/\text{day} = k_3/\text{day}$.

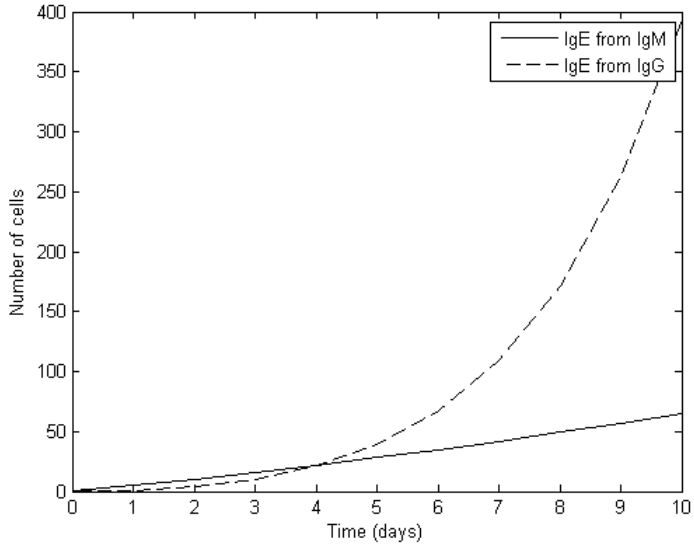

Figure 7: **The variation in the IgE population from IgM and IgG over a 10 day period as predicted by equations (2)-(4).** Here $\alpha_M = \alpha_G = \alpha_E = 0.5/\text{day}$, $d_M = d_G = d_E = 0.1/\text{day}$, $k_1 = 0.3/\text{day}$ and $k_2 = 0.05/\text{day} = k_3/\text{day}$. In this case the majority of cells coming to IgE are from the IgG state.

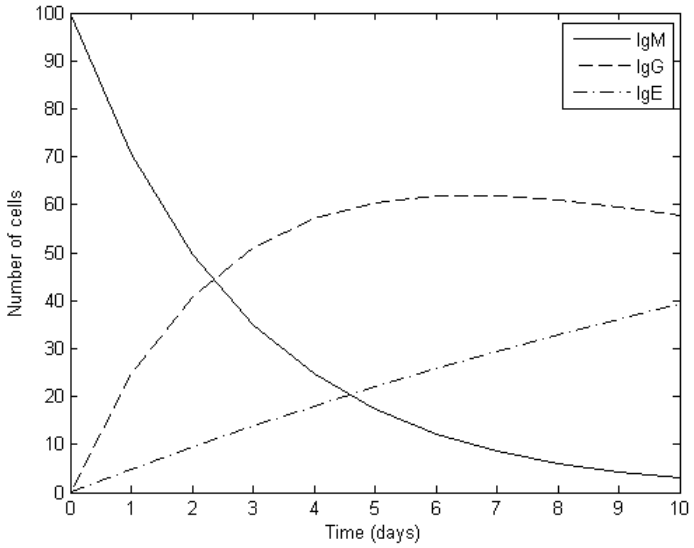

Figure 8: **The variation in IgM, IgG and IgE populations over a 10 day period as predicted by equations (2)-(4) when each cell state has equal rates of proliferation and death.** Here $\alpha_M = \alpha_G = \alpha_E = 0.5/\text{day}$, $d_M = d_G = d_E = 0.5/\text{day}$, $k_1 = 0.3/\text{day}$ and $k_2 = 0.05/\text{day} = k_3/\text{day}$.

IgE populations. As the transition to IgE from IgM is much slower, these cells dominate the IgE populations less.

Whilst various qualitative behaviours in each cell population can be obtained from this model for varying parameter values, with in this case some informed by experimental data (cell death), a range of different transit behaviour is possible. Thus it is imperative in the next stage of determining whether the hypotheses considered here of IgM to IgG transiting more rapidly than IgM to IgG to IgE are correct or need to be considered against others, that model results are compared with other experimental data.

3.4 A birth-and-death process

In [10], the authors suggested an individual-based model of CSR, which takes into account shift dependence on the number of divisions. They simulate the model in discrete time; at each tick of the clock (12 hours) cells are interrogated and their fate is decided according to the following scheme:

The cell object carries information about the number of divisions a cell has undergone and about the type of antibody it is secreting. The advantages of this approach are that the transition rates can be made arbitrarily

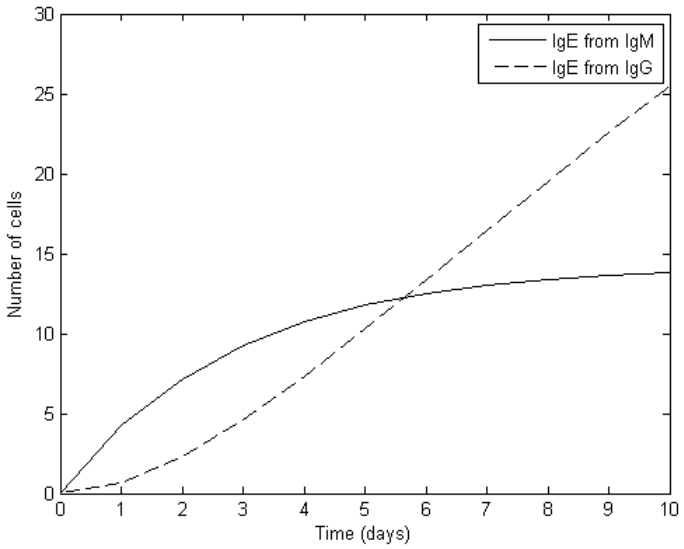

Figure 9: **The variation in the IgE population from IgM and IgG over a 10 day period as predicted by equations (2)-(4) when each cell state has equal rates of proliferation and death.** Here $\alpha_M = \alpha_G = \alpha_E = 0.5/\text{day}$, $d_M = d_G = d_E = 0.5/\text{day}$, $k_1 = 0.3/\text{day}$ and $k_2 = 0.05/\text{day} = k_3/\text{day}$. As in the case where cell proliferation is much greater than that of death the majority of cells coming to IgE are from the IgG state.

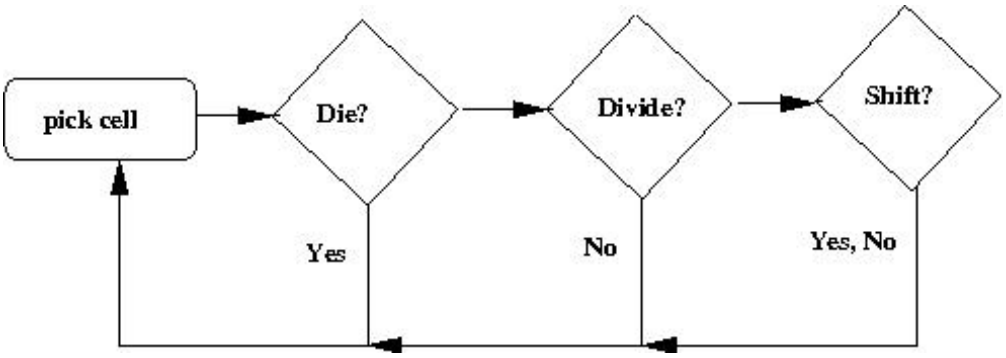

Figure 10: **The scheme of Yaish and Mehr**

complicated.

We can base a simple birth and death process on this scheme; we assume that birth, death, and class shifts are statistically independent processes. The simplest implementation of this idea gives reasonably manageable formulae. If every tick of the clock is 12 hours, and the initial population is N_0, then if P_b and P_d are, respectively, division and apoptosis probabilities, then the number of cells surviving after m ticks is

$$N_m = [(1 - P_d)(1 + P_b)]^m N_0,$$

while the number of cells surviving after m ticks having undergone k divisions is

$$N_{k,m} = \left(\begin{array}{c} m \\ k \end{array} \right) 2^k P_b^k (1 - P_b)^{m-k} (1 - P_d)^m N_0.$$

Computing $N_{k,m}/N_m$ shows that it is actually independent of P_d and has the right behaviour (we take the probability P_b of division per cell between two 12 hour ticks to be $P_b = 0.15$):

The number of cells by tick m that have converted to IgG production is given by the formula

$$N_{m,G} = P_{MG} \sum_{k=1}^{m} N_{k,m} \sum_{l=1}^{k} [(1 - P_{MG})(1 - P_{ME})]^{l-1} (1 - P_{GE})^{k-l},$$

if we assume that the transition probabilities from IgM to IgG, P_{MG}, from IgM to IgE, P_{ME}, and from IgG to IgE, P_{GE}, are all constant. Such a theory gives wrong predictions : the curves of $N_{m,G}/N_m$ versus m are **concave**, while the experimental data gives rise to **convex** ones, but the approach is sufficiently flexible to allow improvements.

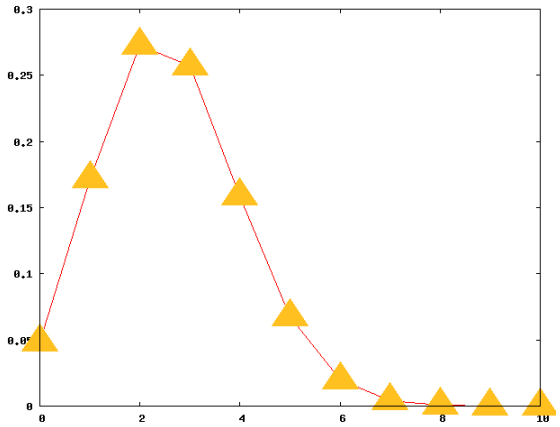

Figure 11: **Proportion of B cells having passed through k divisions versus k after 120 hours.**

So if we assume that all these probabilities are functions of the division number m, e.g.

$$P_{MG} = H(m-2)\alpha(1-\exp(-\beta m)),$$

where $H(\cdot)$ is the Heaviside function, and α and β are positive constants, and similarly for P_{ME} and P_{GE} we have that the probability $\pi_{MG}(k)$ of having performed k divisions, shifting from IgM to IgG production, and staying there is

$$\pi_{MG}(k) = \sum_{l=1}^{k} \left(\prod_{s=1}^{l-1} (1-P_{MG}(s))(1-P_{ME}(s)) \times \right.$$
$$\left. P_{MG}(l) \prod_{r=l+1}^{k} (1-P_{GE}(r)) \right).$$

Then the number of IgG-producing B-cells by the m-th division is given by (I am using the same notation as for constant probabilities)

$$N_{m,G} = \sum_{k=1,m} N_{m,k}\pi_{MG}(k),$$

and thus the proportion we are after is $N_{m,G}/N_m$.

Similarly, the probability $\pi_{ME}(k)$ of having performed k divisions, shifting from IgM to IgE production either directly or via IgG production, is

$$\pi_{ME}(k) = \sum_{l=1}^{k} \prod_{s=1}^{l-1} (1-P_{MG}(s))(1-P_{ME}(s)) \times$$
$$\left(P_{ME}(l) + P_{MG}(l) \sum_{j=l}^{k-1} \prod_{r=l}^{j} (1-P_{GE}(r)) P_{GE}(j+1) \right).$$

Then the number of IgE-producing B-cells by the m-th division is given by

$$N_{m,E} = \sum_{k=1,m} N_{m,k}\pi_{ME}(k).$$

and thus the proportion we after is $N_{m,E}/N_m$.

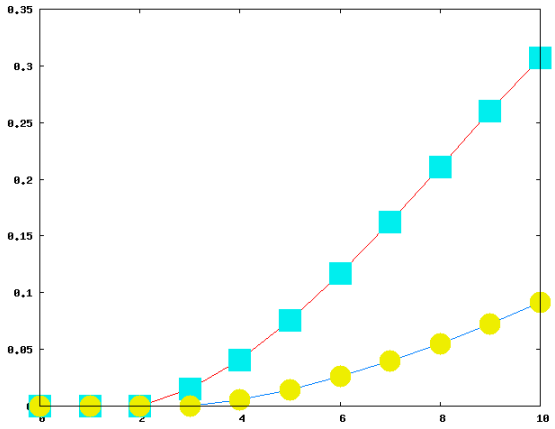

Figure 12: **Proportions of IgG (blue) and IgE (yellow) producing cells versus number of divisions**

In Figure 12 we assume that

$$P_{MG}(l) = 0.2\,H(l-2)(1-\exp(-0.05l));\ P_{ME} = P_{GE} = 0.05\,H(l-3)(1-\exp(-0.05l)).$$

4 Discussion and Conclusions

4.1 Biological viewpoint

To date, while much is known about the cellular dynamics of the germinal centre reactions we still poorly understand the molecular mechanisms that co-ordinate these processes. Several "key" transcription factors have been identified that maintain cellular identity and regulate cell differentiation to plasma cells namely BCL6, BLIMP, PAX5, XBP1, BATF and IRF4 [13, 14, 15]. However, we still know little about the mechanisms that regulate and co-ordinate the early processes of the GC reactions and how these function to control the dynamics of the class switching process.

We proposed to utilise mathematical modeling to help us to identify critical aspects of cellular behavior that might influence the outcome of class switching and combine this with transcriptomic investigations

to identify potential regulators of these activities. The mathematical modeling exercises highlighted the importance of cellular flux (division versus apoptosis) in determining the outcome of the class switching events and matching these to experimental data. In addition, it was found that some element of "division counting" whereby the frequency of switching, apoptosis or proliferation changed as the cells divide was critical to obtaining data from the models that accurately matched experimental results. These findings have not only led the interpretation of the transcriptomic analyses (see later) but have also shaped our future plans to continue this work. We are now in the process of investigating the changes in the rates of apoptosis, cell cycle progression and class switching as a function of division number. Preliminary data from these studies has identified a block in cell cycle progression that appears to occur at the G2 phase of the cell cycle, resulting in an accumulation of cells at this point that is not relieved until later time points in the culture. These data are consistent with our transcriptomic studies that identify a number of known cell cycle checkpoint pathways operating in these early time points. We now plan to continue these studies, allowing us to significantly refine our current mathematical models and further interrogate the key aspects of cellular behavior that control the germinal centre activities.

Further to these studies, initial discussions with Varun Manhas about his finite element PDE solver highlighted the importance of our observations that significant cell movement and accretion occur during our cultures. This conclusion is supported by our transcriptomic analyses that identified chemokine genes (CCL22 and CCL17) as being significantly up-regulated during the cultures. We have now undertaken studies to investigate the rôle, and potential regulators, of these phenomena and their wider importance in determining the outcome of cell fate.

Our investigations of the transcriptomic datasets have identified a number of genes of significant interest in both the regulation of cell flux and behavior in our cultures as well as specific aspects of class switching. Specifically: FGD3 is a Rho GTP exchange factor that is known to regulate the actin cytoskeleton and cell motility [16]; KIF18A is a microtubule motor that is involved in cell migration; while MIIP may inhibit cell migration and regulate mitotic progression. These proteins warrant further investigation for their rôle in regulating the cell migration and accretion phenomenon as well as the cell cycle blockade discussed above. Similarly, CCL17 and CCL22 are potentially very interesting targets; both are known to be powerful Th2 chemoattractants and act to mediate T:B cell interactions during the germinal centre response (reviewed by [17]). However, the potential rôle of these chemokines in B cell chemotaxis has not been previously documented. In support of our hypothesis that these molecules may mediate these activities, is the finding

that their receptor (CCR4) is also up-regulated in these B cells.

NFIL3 has recently been investigated in great depth and found to be central to the regulation of class switching to IgE [18]. We believe the identification of this transcript by our analyses validates the utility of our protocol and likelihood of finding novel regulatory factors.

More specifically focused towards the regulation of class switching to IgE, IKZF2 (Helios) belongs to a family of transcription factors (Ikaros family) that have previously been implicated in the regulation of class switching to IgE [19] and are well known to control lymphoid development [20]. We also believe that Ikaros factors may play an important rôle in directly targeting AID (Activation Induced cytidine Deaminase) to the Ig locus (manuscript under review).

BATF has recently sparked intense interest for its rôle in coordinating class switching [15, 21] However, very little work has been done on BATF3 with regards its B cell functions and it has not been previously studied with regards regulating class switching. We now propose to follow up these studies with more in-depth validation of the expression of the targets accompanied by siRNA mediated interference of their expression to establish their rôle in the regulation of IgE expression.

4.2 Modelling viewpoint

1. The individual-based and birth-death models indicate that assuming constant division and apoptosis rates is sufficient to reproduce the experimental division-number distributions. The birth-death model also allows easy computation of apoptosis and division rates.

2. All our approaches also show that the kinetics of early-stage shift to IgG (and so perhaps also to IgE) is not well described by models that assume constant shift rates; this confirms the findings in the literature that shift rates are division-number dependent.

3. Many of our approaches are flexible enough to incorporate time-inhomogeneity and spatial imhomogeneity.

4. The modelling framework of Yaish and Mehr can be modified: as it is, it ties up division and shifting, but in reality failure to divide is indicative of a checkpoint problem and so could be tied up with death. That the processes are not independent raises interesting challenges for the Gillespie approach and for the birth-death

one.

4.3 Further work

The analyses above have emphasised the importance of "cell flux" in determining the outcome of our B cell cultures. To improve our modeling of these events, it is clear we need to understand better the dynamics of cell division and apoptosis as the cultures progress, and crucially as the cells undergo an increasing number of divisions. To that end, we are now in the process of following apoptosis and cell cycle progression as a function of cell division using the tracking dye CFSE (as utilised in [6]). We believe a better understanding of these dynamics will allow the design of better mathematical models, which will in turn identify and allow us to test other aspects regulating the process of class switching.

In addition, we have initiated preliminary studies to investigate the requirement of the cell clumping phenomenon for these activities: By visualising the aggregation of cells microscopically over time (in conjunction with the studies described above), it is evident that this process begins before cell proliferation and must be dependent upon active cell movement. Preliminary data suggests that if cell aggregation is prevented (by intermittent mechanical disruption of clumps), then cell cycle progression (and in turn division) is retarded. We now intend to investigate this phenomenon in more detail using more sophisticated (and less damaging) methods to inhibit aggregation such as culturing cells in suspension using orbiting platforms. As described, the outcome of these manipulations on cell cycle progression, cell division, apoptosis and class switching will then be investigated.

Once mechanical methods to disrupt aggregation have been investigated, we intend to study the molecular mechanisms driving this event. The bioinformatic analysis of microarray data described, identified two chemokines known to be involved in mediating T:B cell interactions CCL17 and CCL22. We propose to utilise specific, commercially available, antibodies to block the activity of these two chemokines, and their putative receptor (CCR4, described above) to investigate their rôle in the aggregation process. In addition the bioinformatics investigation identified MIIP, a gene thought to regulate cell migration and mitosis, as being significantly differentially expressed. Future studies will investigate the expression of this gene at both the RNA and protein level in these cultures, particularly following blocking of aggregation and "knock-down" of its expression by siRNA methodologies. Similar studies are planned for the other genes of interest described above to begin to determine their rôle in regulating these key aspects of B cell biology.

5 Impact on NC3R Agenda

To date, nearly all the research in this area has been carried out in mice, utilising *in vitro* stimulation of primary murine B cells (isolated post–mortem from the spleen and bone marrow) and mouse models of asthma and allergic disease. These studies use large numbers of animals: typically 3–10 mice per *in vitro* experiment and up to 50 for a full *in vivo* disease model study (particularly if structural lung changes are investigated). These studies are associated with animal welfare concerns; asthma disease models require repeated dosing of inhaled allergen over the course of several weeks and long term kinetic investigations of circulating and lung lavage cell numbers over the course of months in mice with impaired life expectancy. However, these mouse models of disease are far from perfect, mice dont get asthma, while we (and others) have shown that class switching appears to be regulated differently in mice versus man [22]. Our simple human *ex vivo* culture system lends itself easily to manipulation and downstream analysis of consequences and mechanisms. We wish to extend its utility and demonstrate its potential to investigate complex regulatory networks. The development of mathematical models during the study group has already highlighted some key aspects of B cell biology that appear to affect the process of class switching. The importance of these aspects had not been previously recognised in studies of the mouse which show different kinetics of B cell division and many other key aspects of IgE biology from the human system. Further, in combination with our microarray studies we have the ability to identify and manipulate the expression of potential regulatory components and investigate their rôle in controlling IgE production.

Publication of these studies will highlight the utility of this model system both to identify novel regulators and to test potential therapeutics that could not otherwise be studied using murine models. We hope that this would not only contribute to the understanding of asthma and allergy, but also demonstrate the wider application of our system to the general immunobiology community, encouraging the uptake of human *ex vivo* tissue models and reducing current reliance on animal experimentation.

References

[1] Stavnezer, J., Guikema, J. E., and Schrader, C. E. Mechanism and regulation of class switch recombination, Annual Reviw of Immunology (2008), 261–292.

[2] Maul, R. W. and Gearhart, P. J. AID and somatic hypermutation, Advances in Immunology **105** (2010), 159–191 .

[3] R. Küppers, B cells under influence: transformation of B cells by EpsteinBarr virus, Nature Reviews Immunology **3** 92003) 801–812.

[4] Akdis, M. and Akdis, C. A. IgE class switching and cellular memory, Nature Immunology **13** (2012), 312–314 .

[5] L. C. and Zarrin, A. A. The production and regulation of IgE by the immune system, Nature Reviews Immunology **14** (2014), 247–259.

[6] Tangye, S. G. and Hodgkin, P. D. Divide and conquer: the importance of cell division in regulating B-cell responses, Immunology **112** (2004), 509–520.

[7] Duffy, K. R. *et al*. Activation-induced B cell fates are selected by intracellular stochastic competition, Science **335** (2012), 338–341.

[8] Meyer-Hermann, M. *et al*. A Theory of germinal center B Cell selection, division, and exit, CELREP **2** (2012), 162174.

[9] Zhang, Y. *et al*. Germinal center B cells govern their own fate via antibody feedback, Journal of Experimental Medicine **210** (2013), 457464 .

[10] Yaish, B. and Mehr, R. Models for the dynamics and order of immunoglobulin isotype switching, Bulletin of Mathematical Biology **67** (2005), 15–32.

[11] Kingsford, C. and Salzberg, S. L. What are decision trees? Nature Biotechnology **26** (2008), 1011–1013.

[12] Therneau, T. M. and Atkinson, B. (R port by Brian Ripley), rpart: Recursive Partitioning. Oxford, UK2012 [cited 2014 18/03/2014]; Recursive partitioning for classification, regression and survival trees. An implementation of most of the functionality of the 1984 book by Breiman, Friedman, Olshen and Stone]. Available from: `http://cran.r-project.org/web/packages/rpart/index.html`.

[13] Murphy, T. L, Tussiwand, R. and Murphy, K. M. Specificity through cooperation: BATFIRF interactions control immune-regulatory networks, Nature Reviews Immunology **13** (2013), 499-509.

[14] Igarashi, K. *et al*. Orchestration of plasma cell differentiation by Bach2 and its gene regulatory network, Immunological Reviews **261** (2014), 116-125.

[15] Ise, W. *et al*. The transcription factor BATF controls the global regulators of class-switch recombination in both B cells and T cells, Nature Immunology **12** (2011), 536–543.

[16] Hayakawa M. *et al.* Novel insights into FGD3, a putative GEF for Cdc42, that undergoes $SCF^{FWD1/\beta TrCP}$-mediated proteasomal degradation analogous to that of its homologue FGD1 but regulates cell morphology and motility differently from FGD1, Genes to Cells **13** (2008), 329–342.

[17] Holgate, S. T. Innate and adaptive immune responses in asthma, Nature Medicine **18** (2012), 673–638.

[18] Kashiwada, M. *et al*. IL-4-induced transcription factor NFIL3/E4BP4 controls IgE class switching, Proceedings of the National Academy of Sciences USA **107** (2010), 821–826.

[19] Sellars, M., Reina-San-Martin, B., Kastner, P. and Chan, S. Ikaros controls isotype selection during immunoglobulin class switch recombination, The Journal of Experimental Medicine **206**(2009), 1073–1087.

[20] Schwickert, T. A. et al. Stage-specific control of early B cell development by the transcription factor Ikaros, Nature Immunology **15** (2014), 283–293.

[21] Zan, H. and Casali, P. Regulation of *Aicda* expression and AID activity, Autoimmunity **46** (2013), 83–101.

[22] Gould, H. J. and Ramadani, F. IgE Responses in mouse and man and the persistence of IgE memory, Trends in Immunology, in press.