**Thesis submitted for the degree of Doctor of Philosophy**

# Impact of genetic variation on gene regulatory effects of vitamin D in immunity and inflammation

**Antonio Jorge Berlanga-Taylor**
**MBBS, MSc**

**Wolfson College**
**University of Oxford**
**Trinity 2013**

## Supervisors

Dr. Julian C. Knight

Wellcome Trust Centre for Human Genetics

Nuffield Department of Clinical Medicine

Medical Sciences Division

University of Oxford

Prof. George C. Ebers

Wellcome Trust Centre for Human Genetics

Nuffield Department of Clinical Neurosciences

Medical Sciences Division

University of Oxford

# Dedication

To my parents, Sue and Antonio, and my sister, Jessica, whose love, laughter and teachings have always told me where to go.

To many friends and family for their kindness and unwavering support, if I cannot name them all at least some who represent all those close to me who have touched my life and helped me along the way: Claudia and Ricardo, the Taylors, Carlos and Maite, Julia and Manuel, los del kinder, la banda Puma...

To Zahra, who has made the last two years some of the best in my life and has withstood the side effects of being with a DPhil candidate.

I have learnt from many others and I am grateful to them.

# Acknowledgements

This thesis is the product of work I carried out at the Wellcome Trust Centre for Human Genetics, University of Oxford, from 2009 to 2013 under the supervision of Dr. Julian C. Knight and Prof. George C. Ebers. I led the process behind each chapter but the research would not have been at all possible without the help and intervention of many people. First and foremost I am deeply indebted to my supervisors for providing clarity, direction, resources and a highly supportive environment to work in. I would also like to thank the post-doctoral fellows in both labs for their continued support, particularly Dr. Julia Morahan, for her guidance and patience. I wish to thank the members of the Knight Lab and of the Ebers Lab for their help and friendship throughout these years. I would also like to thank the research nurses and principal investigators of the Genomic Advances in Sepsis (GAinS) study, particularly Prof. Charles Hinds, for access and support while I was carrying out research in sepsis. I would like to express great gratitude to Dr. Jayachandran Radhakrishnan for providing access to the GAinS gene expression data and to Dr. Peter Humburg for invaluable help and guidance with some of the computational aspects of this work.

The work in this thesis was supported by DPhil studentships from the Multiple Sclerosis Society (MSS, UK) and the National Council for Science and Technology (CONACyT, Mexico). Their financial assistance was essential.

None of the work presented here would have been possible without the trust, faith and altruism of the many anonymous patients and their families who donated samples for

clinical research. I would like to express my deep respect and appreciation for them. Many are no longer with us.

# Attributions

When I started the DPhil a genome-wide study to map the vitamin D receptor (VDR) binding sites was on-going (Ramagopalan et al., 2010). This was a collaborative effort between the Ebers Lab, Dr. Andreas Heger (CGAT, MRC FGU, University of Oxford), Prof. Chris Ponting (MRC FGU, University of Oxford) and Dr. Julian Knight. My participation was restricted to assisting in data analysis regarding disease associated variants and VDR binding sites, data interpretation and participating in the writing of the manuscript.

The study to investigate the role of genetic variants in VDR binding using in vitro techniques was conceived and designed by Dr. Julian Knight, Dr. Julia Morahan (Ebers Lab) and I. The experiments, analysis and interpretation were my work with help from Dr. Julian Knight, Ms. Katie Morrison (Ebers Lab) and Dr. Julia Morahan.

The family based test of association was carried out in patients and family members from the Canadian Collaborative Project on the Genetic Susceptibility to Multiple Sclerosis (CCPGSMS). This is an on-going cohort led by Prof. George Ebers that began over 30 years ago with the participation of multiple labs and clinics. DNA samples have been extracted and prepared over the years and a common resource for multiplex genotyping was available at the Ebers Lab when I joined. Dr. Sreeram Ramagopalan and I conceived and designed the study presented here. Data analysis was carried out by Ms. Andrea Para and Dr. Mathew Lincoln (former members of the Ebers Lab). Dr. Sreeram Ramagopalan and I interpreted and contextualised the results.

The volunteer cohort project was conceived by Dr. Julian Knight, led by him and co-led by Dr. Benjamin Fairfax (J Knight Lab). The recruitment of volunteers, experimental procedures and data analysis concerning genetic variants and gene expression were done by Dr. Benjamin Fairfax, Ms. Seiko Makino (J Knight Lab) and Dr. Katharine Plant (J Knight Lab). My participation consisted of conceiving, designing and leading a genotype-transcriptome-plasma study and plasma samples were thus kept. Vitamin D in plasma was measured by the Homerton Hospital Clinical Laboratory at my request. I designed the analytical strategy for this study (vitamin D in healthy volunteers). Dr. Jayachandran Radhakrishnan (J Knight Lab) and I carried out the analysis and I am responsible for interpretation and writing.

Prof. Charles Hinds of Queen Mary University of London and Dr. Chris Garrard of the John Radcliffe Hospital, Oxford, conceived, designed and initiated the Genomic Advances in Sepsis (GAinS) study as chief investigators of the UK Critical Care Genomics (UKCCG) group (www.ukccg-gains.org/index.htm). This group aims to understand the role of genetics in sepsis outcome and is a national multi-centre observational study with long-term follow-up for mortality. It has used candidate gene and genome-wide approaches to understand genetic associations to severe sepsis and septic shock. An increased focus on functional genomics was initiated with the participation of Dr. Julian Knight. Several phases of GAinS were on-going or had ended when I arrived. My participation consisted in conceiving and leading the vitamin D in sepsis study through GAinS. The recruitment of patients was carried out by Paula Hutton (GAinS nurse coordinator) and other GAinS research nurses and investigators. Drs. Jayachandran Radhakrishnan (J. Knight Lab), Narelle Magueri (J.

Knight Lab), Eduardo Svoren (Hinds Lab) and Ms. Emma Davenport (J Knight Lab) performed the RNA extractions for GAinS. Microarray hybridisation was done by Mr. Peter Ellis at the Sanger Institute, Cambridge, at the request of Dr. Julian Knight and Dr. Jayachandran Radhakrishnan. Data and samples were managed by Ms. Emma Davenport. I conceived the hypothesis, measured vitamin D in plasma through the Homerton Hospital laboratory, measured cathelicidin in plasma and analysed the epidemiological data. I designed the analytical strategy for the molecular associations and performed the data analyses. Dr. Peter Humburg (J. Knight Lab) analysed the gene expression data presented and also advised and helped me at several stages of the vitamin D in GAinS study. I interpreted the data and wrote the manuscript with Dr. Julian Knight and Prof. Charles Hinds.

Finally, I am indebted to all the members of the J Knight and Ebers labs for multiple helpful discussions during the course of my work.

# Publications

Berlanga-Taylor AJ, Disanto G, Ebers GC, Ramagopalan SV. Vitamin D-gene interactions in multiple sclerosis. J Neurol Sci. 2011 Dec 15;311(1-2):32-6.

Berlanga-Taylor AJ, Disanto G, Ramagopalan SV. Re: "Use of penicillin and other antibiotics and risk of multiple sclerosis: a population-based case-control study". Am J Epidemiol. 2012 Apr 1;175(7):727-8; author reply 8-9.

Berlanga-Taylor AJ, Ramagopalan SV. Vitamin D and multiple sclerosis: what is the clinical impact? Expert Opin Med Diagn. 2013 Apr 9.

Disanto G, Berlanga AJ, Handel AE, Para AE, Burrell AM, Fries A, Handunnetthi L, et al. Heterogeneity in multiple sclerosis: scratching the surface of a complex disease. Autoimmune Dis. 2010;2011:932351.

Disanto G, Hall C, Lucas R, Ponsonby AL, Berlanga-Taylor AJ, Giovannoni G, Ramagopalan SV. Assessing interactions between HLA-DRB1*15 and infectious mononucleosis on the risk of multiple sclerosis. Mult Scler. 2013 Feb 14.

Disanto G, Sandve GK, Berlanga-Taylor AJ, Morahan JM, Dobson R, Giovannoni G, Ramagopalan SV. Genomic regions associated with multiple sclerosis are active in B cells. PLoS One. 2012;7(3):e32281.

Disanto G, Sandve GK, Berlanga-Taylor AJ, Ragnedda G, Morahan JM, Watson CT, Giovannoni G, et al. Vitamin D receptor binding, chromatin states and association with multiple sclerosis. Hum Mol Genet. 2012 Aug 15;21(16):3575-86.

Handel AE, Handunnetthi L, Berlanga AJ, Watson CT, Morahan JM, Ramagopalan SV. The effect of single nucleotide polymorphisms from genome-wide association studies in multiple sclerosis on gene expression. PLoS One. 2010;5(4):e10142.

Ramagopalan SV, Dyment DA, Cader MZ, Morrison KM, Disanto G, Morahan JM, Berlanga-Taylor AJ, et al. Rare variants in the CYP27B1 gene are associated with multiple sclerosis. Ann Neurol. 2011 Dec;70(6):881-6.

Ramagopalan SV, Heger A, Berlanga AJ, Maugeri NJ, Lincoln MR, Burrell A, Handunnetthi L, et al. A ChIP-seq defined genome-wide map of vitamin D receptor binding: associations with disease and evolution. Genome Res. 2010 Oct;20(10):1352-60.

## Funding sources

I received scholarships from the National Council for Science and Technology (CONACyT, Mexico) and the Multiple Sclerosis Society of Great Britain and Northern Ireland and had access to research funds from the Multiple Sclerosis Society of Canada Scientific Research Foundation.

# Abbreviations

$1,25(OH)_2D$ = 1,25-dihydroxyvitamin D; calcitriol

25(OH)D = 25-hydroxyvitamin D; calcifediol

Cholecalciferol = vitamin $D_3$

Ergocalciferol = vitamin $D_2$

APC = antigen presenting cell

CAP = community acquired pneumonia

CD = coeliac disease

CI = confidence interval

CNV = copy number variation

FDR = false discovery rate

FP = faecal peritonitis

GAinS = genomic advances in sepsis

GWAS = genome-wide association study

HLA = human leukocyte antigen

HMG = high mobility group

ICU = intensive care unit

IL = interleukin

LD = linkage disequilibrium

MHC = major histocompatibility complex

MS = multiple sclerosis

PMA = phorbol 12-myristate 13-acetate

RA = rheumatoid arthritis

RXR = retinoid X receptor

SNP = single nucleotide polymorphism

T1D = type 1 diabetes

Th = T helper cell

TLRs = toll like receptors

VDR = vitamin D receptor

VDRE = vitamin D response element

# ABSTRACT

Genome-wide association studies in multifactorial diseases have contributed to our understanding of genetic risk and defined specific disease-associated loci in particular populations. However, risk cannot be fully explained by genetics and evidence points to both genetic and environmental factors being important in causation and pathophysiology. The role of vitamin D in calcium homeostasis is well established. Over the last 30 years it has become clear that vitamin D has a wider role in physiology and disease, notably in autoimmune, cancer and infectious conditions. However, the molecular mechanisms and possible causal role of these associations are poorly understood. Here I propose that the role of vitamin D in immune and inflammatory responses is significant, that genetic variation partly determines the response to vitamin D and that integrative analysis can yield important insights for disease mechanisms. For this I investigate the relationship between vitamin D and genetic risk involving the immune system by focusing on multiple sclerosis and sepsis, conditions classically defined as autoimmune and inflammatory respectively. I describe data resolving genetic variation associated with autoimmune diseases in vitamin D receptor binding sites; the association to multiple sclerosis of a genetic variant lying within a VDR binding site; the correlation of plasma vitamin D with genotype and cell specific transcriptomes in healthy volunteers; and the extent of vitamin D deficiency in severe sepsis and septic shock, its association with survival, correlation with gene expression and use in sub-classification to identify patients at higher risk of death. The limitations of each study and future work are discussed. Integrating epidemiological and clinical observations with genetic and functional

genomics techniques has the potential to reveal interactions in population specific disease susceptibility that may lead to an improved understanding of disease mechanisms and clinical translation. The work I present here bridges molecular analysis, candidate and genome- wide, with phenotypic observations that are important in our understanding of disease.

# Table of contents

# Detailed table of contents

# List of Tables

# List of Figures

# CHAPTER 1: INTRODUCTION

## General Thesis Structure

### SUMMARY AND RATIONALE

The work presented here is based on the premise that vitamin D is likely to impact on disease processes, particularly those of autoimmunity (exemplified by multiple sclerosis) and inflammatory disease (exemplified by sepsis). The overarching thesis is that vitamin D modulates immunity and that its regulation, blood levels and actions are important in disease. The hypotheses follow from previous work carried out in the Ebers and Knight Labs defining the vitamin D receptor (VDR) binding map in B cells (Ramagopalan et al., 2010; Ramagopalan et al., 2009). I address this thesis using a combination of clinical epidemiological approaches with functional genomics and molecular genetics methods. Below I explain in detail the thesis I put forward as well as the general framework and hypotheses that are tested.

### GENERAL HYPOTHESES

1. ***Molecular characterisation of genetic regulatory variants and vitamin D receptor binding in autoimmune disease***

Published genome-wide analyses in terms of disease association and mapping of protein-DNA interactions by ChIP-seq have highlighted disease associated loci that overlap with VDR binding sites. MS is an interesting paradigm for such work although other autoimmune diseases may also be relevant. The overarching question is thus: 'Is there a relationship between disease associated loci and VDR binding in

immune related diseases?' This leads to the testable hypothesis that disease associated genetic variation disrupts VDR binding. Chapter three investigates this and covers the following specific questions:

- Is VDR binding enriched near disease-associated genomic intervals?
- Are disease associated single nucleotide polymorphisms (SNPs) present within VDR motifs? Do these modulate VDR binding?
- Are genetic variants found within VDR binding intervals near MS-associated genes? Are these variants associated to MS susceptibility?

***2. Integrative analysis of genotype - gene expression - phenotype associations of plasma vitamin D in healthy volunteers and in patients with severe sepsis and septic shock***

Over the last several years it has come to light that vitamin D plays an important role in immune and anti-inflammatory responses. The contribution of genetic variants to modulating the effects of vitamin D remains poorly understood however and the importance of vitamin D in inflammatory diseases has not been fully investigated. Sepsis remains a disease with high mortality, heterogeneous presentation but common outcome and lacks targeted therapeutic options. Studies performing integrated analyses taking into account cell specificity, genotypes and/or transcriptomes in healthy or diseased individuals are generally scarce but may provide important insight. In this section I investigate the relevance of vitamin D to sepsis and if an integrative analysis of phenotype, vitamin D and additional biological layers can yield

better understanding of vitamin D biology and disease mechanisms. The hypotheses tested in chapters four and five are: i) genotype-transcriptome-plasma 25-hydroxyvitamin D are associated to each other and are cell type specific in healthy individuals; ii) patients with severe sepsis or septic shock are vitamin D deficient, which is associated to outcome and integrative analysis can aid in disease classification.

The specific questions covered in chapter four are:

- What is the distribution of plasma levels of vitamin D in healthy individuals?
- Can I replicate GWAS findings of variants associated to vitamin D deficiency using a candidate approach in a cohort of healthy individuals?
- Is there a relationship between gene expression and plasma levels of 25-hydroxyvitamin D in healthy volunteers? Is this relationship cell specific?
- Does pathway analysis reveal informative networks relevant for disease associated research?

The specific questions covered in chapter five are:

- Are patients with severe sepsis or septic shock vitamin D deficient? If so, is this associated with survival? Can vitamin D levels help predict sepsis outcome?
- Can integrating gene expression, plasma vitamin D levels and relevant clinical and molecular information help identify sub-groups of patients?

## AIMS AND OBJECTIVES

The effect of genetic variation and the molecular mechanisms underlying vitamin D regulatory effects remain incompletely understood. The overall objectives of this work are to characterise the impact of genetic variation on regulation exerted by vitamin D and evaluate the importance of vitamin D on disease within the context of inflammation and immunity. The specific aims include:

- Understand the role of vitamin D receptor binding and disease associated variants:
    - To determine the enrichment of VDR binding near disease associated genomic intervals;
    - To localise disease associated genetic variation which may affect vitamin D receptor binding;
    - To determine the effect of disease-associated genetic variation on VDR binding;
    - To determine the association of genetic variation present within VDR binding intervals and MS susceptibility.
- Understand the role of vitamin D in a cohort of healthy individuals:
    - To replicate GWAS findings of variants associated to vitamin D deficiency using a candidate approach in a cohort of healthy individuals;

- To analyse the relationship between cell-specific gene expression and serum 25-hydroxyvitamin $D_3$ in healthy individuals;
- To interpret integrated analysis findings in the context of disease relevance and pathway analysis.

- Understand the role of vitamin D in patients with severe sepsis and septic shock:
  - To analyse the distribution of plasma 25-hydroxyvitamin $D_3$;
  - To analyse the correlation of plasma vitamin D deficiency to outcome and possible confounding factors;
  - To analyse the relationship between leukocyte gene expression, plasma 25-hydroxyvitamin $D_3$ and other sepsis outcome risk factors;
  - To test whether an integrated analysis using molecular and clinical information can identify sub-groups of patients.

**THESIS ORGANISATION**

The thesis covers various aspects of vitamin D biology and disease and I use several experimental and computational approaches. As such I have provided relevant literature reviews, introductions and discussions for each section to fully explain the rationale and results behind each study. The manuscript is structured in the following way:

Chapter one provides a review that introduces vitamin D biology and disease. I use two conditions, MS and sepsis, to frame the analysis of vitamin D actions and investigate its importance in disease. Chapter two details the materials and methods used in all experiments. Chapters three, four and five contain the experimental data, results and interpretation of the studies I performed.

Chapter three describes genetic variation associated to disease and vitamin D receptor binding. In this chapter I first use molecular biology and bioinformatics techniques to understand the impact of autoimmune disease associated genetic variation on VDR binding. Additionally, I explore the association of vitamin D receptor binding, genetic variation and susceptibility to MS.

Chapters four and five use an integrated analytical approach based on observations in two cohorts, one healthy and one with severe sepsis and septic shock, with the aim of understanding the relationship between genotype, gene expression and plasma vitamin D in the context of inflammation.

In chapter four I explore the *in vivo* relationship of plasma vitamin D to genotype and to cell specific gene expression in healthy individuals. In chapter five I explore the importance of vitamin D deficiency in sepsis, show how clinical observations of survival risk factors are mirrored at the molecular level and how these can be used to sub-classify patients.

Chapter six is a general discussion that integrates the different findings and contextualises them under the broader theme of vitamin D research in immunity and inflammation. It also assesses the strength of evidence regarding causality of vitamin

D deficiency in critical illness and sepsis. Additional results of each experimental section are included as appendices for further reference.

# Literature review

In the 1900s vitamin D deficiency was found to be the cause of rickets and for many decades it was considered that vitamin D was solely for optimal bone health. However, there is now considerable evidence that vitamin D deficiency underlies risk for several diseases including autoimmune and inflammatory conditions (Holick, 2007). It has been estimated that over one billion people worldwide lack vitamin D to some degree (Holick, 2007) due to dietary deficiency or inadequate sun exposure.

Vitamin D supplements have been associated with reduced risk for several diseases, including MS, rheumatoid arthritis, and type 1 diabetes (Ebers, 2008). Importantly, a meta-analysis of randomised controlled studies found a statistically significant 7% reduction in mortality from any cause in individuals receiving vitamin D supplementation (Autier et al., 2007). Deficiency of vitamin D has also been shown to increase the risk to some of these conditions, notably MS (Munger et al., 2006).

Despite many research efforts, it is still incompletely understood how vitamin D acts at the molecular level in diseased states. Recent technological progress, namely in functional genomics methods through applications of next-generation DNA sequencing to DNA-protein interactions, permit high throughput studies that attempt to answer these questions with greater depth and accuracy than before (Alekseyenko et al., 2008; Park, 2009).

Here I present work that provides insight into the actions of vitamin D in disease. I focus on immune and inflammatory mechanisms using MS and sepsis as examples. In

this literature review I first discuss vitamin D biology and then provide separate accounts of the current general and vitamin D related knowledge of MS and the sepsis syndrome.

# Vitamin D biology

The first description of rickets was made in 1645 by Daniel Whistler in Leiden and shortly after in 1650 by Francis Glisson in London (Bouillon et al., 2008). The description of dietary vitamin D and the discovery that UV irradiation of vegetarian oil produced vitamin $D_2$ eventually led to the identification of the origin of rickets (Bouillon et al., 2008; Brumbaugh et al., 1975; Fraser et al., 1970). The cloning of the VDR in several species and the discovery that it is expressed in virtually all human tissues made it clear that its functions were important and varied (McDonnell et al., 1987).

## SOURCES AND METABOLISM OF VITAMIN D

The main source of vitamin D in humans is skin exposure to sunlight. Pre-vitamin $D_3$ is formed in the skin upon exposure of 7-dehydrocholesterol (DHCR7) to ultraviolet B radiation at wavelengths of 290 to 315 nm (Aris et al., 2005). It is then immediately converted to vitamin $D_3$ in a heat dependent process (Holick, 2007). Regardless of its source, vitamin $D_3$ (or vitamin $D_2$ from diet or supplements) is transported by the lymphatics in chylomicrons and can be stored in adipocytes or transported to other tissues through binding to the vitamin D–binding protein (Holick, 2007) (DeLuca, 2004).

Vitamin D is metabolised to 25-hydroxyvitamin D (25(OH)D) by vitamin D-25-hydroxylase (*CYP2R1*) mainly in the liver, though the hydroxylase is expressed in multiple tissues (Horst et al., 2005; Hsu et al., 2006). 25(OH)D is converted to its active form, 1,25-dihydroxyvitamin D ($1,25(OH)_2D$), by 25-hydroxyvitamin D-1α-

hydroxylase (1-OHase; *CYP27B1*), an enzyme highly expressed in the kidneys though now known to be found in multiple tissues (Horst et al., 2005).

Although vitamin D metabolism has been well studied, its regulation is complex and tissue dependent. Several molecules including serum phosphorus, calcium, and fibroblast growth factor 23 control the expression of *CYP27B1*. Negative feedback loops cause 1,25$(OH)_2$D to decrease its own synthesis and that of the parathyroid hormone. Catabolism of 1,25$(OH)_2$D occurs through hydroxylation at the 24th carbon position by 1,25-dihydroxyvitamin D 24-hydroxylase (*CYP24A1*) to generate the bile excreted inactive form calcitroic acid (Horst et al., 2005). Increased expression of CYP24A1 leads to catabolism and thus locally produced 1,25$(OH)_2$D does not enter the circulation and lacks systemic effects (Holick, 2007; Horst et al., 2005).

At the molecular level, vitamin D exerts its influence through VDR though independent actions of both molecules have been described. The *VDR* gene encodes the nuclear hormone receptor for vitamin $D_3$ and upon activation by 1,25$(OH)_2$D it forms a heterodimer with the retinoid-X receptor (RXR), translocates to the nucleus and binds to specific genomic sequences (vitamin D response elements, VDREs) where it regulates transcriptional responses (Pike et al., 2010).

Although the most important source for vitamin D is by exposure to sunlight, individuals at risk require dietary supplementation. This is increasingly true for populations away from the equator, individuals with dark skin, those who use excessive sunscreen and those who do not expose their skin to sunlight for religious

or lifestyle reasons. As such, adequate guidelines for sun exposure and vitamin D supplementation remain necessary, particularly for at risk groups.

### GENETIC DETERMINANTS OF PLASMA VITAMIN D LEVELS

The circulating form of vitamin D, 25(OH)D, is the most commonly measured and is used to describe an individual's status (DeLuca, 2004). The definition of normality for vitamin D status is under debate, particularly for non-skeletal functions. A recent report from the Institute of Medicine in the USA highlighted the need for rigorous evidence to be collected for non-classic effects of vitamin D and recommended current plasma levels to be at 50 to 150 nmol/L (Ross et al., 2011).

Recent studies have indicated that heritability of 25(OH)D in plasma can be high. It is estimated that ~25% of inter-individual variability can be attributed to season of measurement, latitude or intake (Wang et al., 2010). A study in male monozygotic twins determined that there is significant seasonal variation with a 15 nmol/L lower value during the winter than during the summer. It was estimated that 70% of the variation during winter was explained by genetic factors though in the summer concentrations did not have a discernible heritable component (Karohl et al., 2010). Karohl and colleagues (2010) also confirmed a previous finding of significantly higher intra-class correlation between monozygotic twins as compared to dizygotic twins (Orton et al., 2008).

Two recent GWAS found that common variants present in *CYP2R1, 7-DHCR, CYP24A1* and in the vitamin D binding protein (*GC, DBP*), were associated with 25(OH)D blood levels (Wang et al., 2010). These genetic markers and genotype

combinations were associated to increased risk of vitamin D insufficiency (<75 nmol/L), deficiency (<50 nmol/L) and severe deficiency (<20 nmol/L) (Ahn et al., 2010; Wang et al., 2010). Interestingly, rs2282679 within the *GC* gene was also significantly associated to concentrations of GC protein with the minor allele associated to lower levels (Table 1) (Wang et al., 2010).

**Table 1 Single nucleotide polymorphisms associated to vitamin D plasma levels by genome-wide association studies.**

| Reference | Reported Gene(s) | Strongest SNP | Risk Allele Frequency | P-value |
|---|---|---|---|---|
| Ahn et al 2010 | *GC* | rs2282679 | 0.26 | 2.00E-49 |
| | *CYP2R1* | rs2060793 | 0.41 | 3.00E-17 |
| | *NADSYN1,DHCR7* | rs3829251 | 0.19 | 3.00E-09 |
| Wang et al 2010 | *GC* | rs2282679 | 0.29 | 2.00E-109 |
| | *CYP2R1* | rs10741657 | 0.40 | 3.00E-20 |
| | *NADSYN1, DHCR7* | rs12785878 | 0.23 | 2.00E-27 |

### ROLE IN CALCIUM HOMEOSTASIS

Vitamin D is known to play a central role in calcium homeostasis by increasing the availability of calcium and promoting bone cell maturation. It has long been known that a cause of rickets and osteomalacia is vitamin D deficiency. Reports indicate that only ~15% of dietary calcium and ~60% of phosphorus would be absorbed without vitamin D and that serum levels of 25(OH)D are directly related to bone mineral density (Bischoff-Ferrari et al., 2006). Serum 25(OH)D of 100 nmol/L or more are required to reach peak densities while levels of 75 nmol/L or less show decreased intestinal calcium absorption with increased production of parathyroid hormone

(PTH) (Holick, 2007). PTH is central to bone metabolism and forms feedback control loops that increase production of $1,25(OH)_2D$ and serum calcium concentrations through several mechanisms. The role of vitamin D in calcium homeostasis and other systems has been reviewed extensively however and can be referred to elsewhere (Bouillon et al., 2008; Holick, 2007). For the remainder of this manuscript I will focus on the relevance of vitamin D to immunity and inflammation, in particular in reference to MS and sepsis.

**ROLE IN IMMUNE AND INFLAMMATORY PROCESSES**

Ample evidence has been gathered in epidemiological and molecular studies in both humans and in animal models that support the role of vitamin D in immune homeostasis (Figure 1). The discovery that multiple cell types express VDR, including many in immune, breast, prostate, brain and colon tissues and respond to $1,25(OH)_2D$, made it clear that the actions of vitamin D are ubiquitous and pleiotropic (DeLuca, 2004; Holick, 2007). Additionally, many of these tissues also express the metabolic enzymes CYP27B1 and CYP24A1, including B cell subpopulations (Morgan et al., 1999). In monocytes and macrophages vitamin D appears to favour differentiation, enhance their antimicrobial effects by increasing the production of cathelicidin and augment chemotactic and phagocytic responses (Gombart, 2009) (Xu et al., 1993). In myeloid dendritic cells a decrease in critical cell surface receptors and chemokine receptors, including CD40, CD80, CD86, CCR7, IL12, amongst others, as well as an increase in mannose receptors and key interleukins (IL-3, IL-10) and chemokine ligands (CCL2, CCL18, CCL22) following effects of $1,25(OH)_2D$ have been observed (Baeke et al., 2010). These and other effects may impair allo- and self-

reactive T cell activation by inhibiting the differentiation, maturation, activation, and survival of dendritic cells (Penna et al., 2000). Moreover, in vitro pre-conditioning with $1,25(OH)_2D_3$ appears to generate regulatory dendritic cells that selectively induce apoptosis of self-reactive T cells without affecting other T cell populations (van Halteren et al., 2004).

In B cells $1,25(OH)_2D$ seems to have a general repressive effect as it inhibits plasma cell differentiation and proliferation; induces apoptosis of activated B cells and inhibits immunoglobulin secretion (Baeke et al., 2010) (Chen et al., 2007). CCR10 expression promotes mucosal immunity by directing mature B cells and is induced by $1,25(OH)_2D_3$ in terminally differentiating human B cells (Shirakawa et al., 2008). Calcitriol ($1,25(OH)_2D_3$) in activated B cells increases IL-10 production and may provide a pathway for immune reactivity suppression (Fillatreau et al., 2002; Heine et al., 2008; Shirakawa et al., 2008).

In T cells, vitamin D inhibits Th1 and Th17 development and cell proliferation while favouring a Th2 phenotype; it down regulates CD95, IL-2 and IFN-gamma, induces hypo-responsiveness to self-antigens and increases the frequency of regulatory T cells (Baeke et al., 2010; Matilainen et al., 2010; Penna et al., 2005). On the contrary, it has been claimed that vitamin D may activate T cells by inducing phospholipase C-gamma1 though this has recently been brought into question (Smolders et al., 2010; von Essen et al., 2010). Variable results have been observed regarding production of IL-10 and IL-4 and this may be due to different effects on T cell subpopulations and local milieu.

Importantly, recent studies probing the immune response of patients suffering hereditary 1,25$(OH)_2$D–resistant rickets (OMIM 277440; HVDRR) have shown impaired function (Tiosano et al., 2013). HVDRR is a condition caused by genetic mutations that truncate VDR with fewer than 100 patients known (Tiosano et al., 2012). HVDDR serves as a natural experiment of a VDR knockout in humans and can provide insight into disease mechanisms. A subset of 35 patients with HVDDR has been followed for several years and, although it appears that they do not suffer from increased infections or autoimmune conditions, they do show altered immune responses (Tiosano et al., 2013). It was initially shown that macrophages and neutrophils from HVDDR sufferers have defective fungicidal activity (Etzioni et al., 1989) and more recently, using a case-control design with *ex-vivo* analysis, Tiosano and colleagues demonstrated that HVDDR patients have impaired responses in both innate and adaptive components (Tiosano et al., 2013).

This recent study showed that the antimicrobial peptide cathelicidin was expressed at lower levels in monocytes of patients and administration of 25$(OH)D_3$ increased the expression of cathelicidin, VDR, C/EBPϵ , C/EBPβ and hnRNP while suppressing TLR2 in control monocytes only. Interestingly, monocytes from patients showed increased expression of other molecules important in the immune response (CYP27B1, C/EBPα, NOS2, and ARG1) with higher concentrations of pro-inflammatory cytokines (TNF and IL-17) in HVDRR lymphocyte cultures than in those of controls. Furthermore, administration of 25$(OH)D_3$ in culture suppressed IL-17, a pro-inflammatory mediator, only in control lymphocytes while increasing the

expression of anti-inflammatory molecules (IL-4, IL-10, and interferon-γ) in control lymphocyte media but not in that of HVDRR patients (Tiosano et al., 2013).

In general, it appears that the actions of vitamin D in immunity depend on the cell type as well as the stage of differentiation. Other factors, such as type of vitamin D, dose, developmental stage, and environmental input and cell-cell interactions may also be important. Differing vitamin D effects on cytokine production according to microbial presence and season have been observed with $1,25(OH)_2D$ produced in monocytes and macrophages having intracrine and paracrine effects on activated T and B lymphocytes and other surrounding tissues (Khoo et al., 2011; Khoo et al., 2011; Khoo et al., 2011). Robustness in immune and inflammatory responses may compensate partially or fully for alterations in vitamin D physiology as may be the case in HVDRR patients.

On the whole, vitamin D seems to exert many actions on the immune system and at least for autoimmunity it may regulate harmful responses by enhancing dendritic cell control, increasing T and B regulatory cells, reducing T and B cell proliferation and shifting the T cell response to a Th2 phenotype. The complexity of the regulatory mechanism of VDR/RXR-$1,25(OH)_2D_3$ in the immune system is still to be fully understood and it remains an open question as to how it effectively controls gene expression across different cell types and in different conditions.

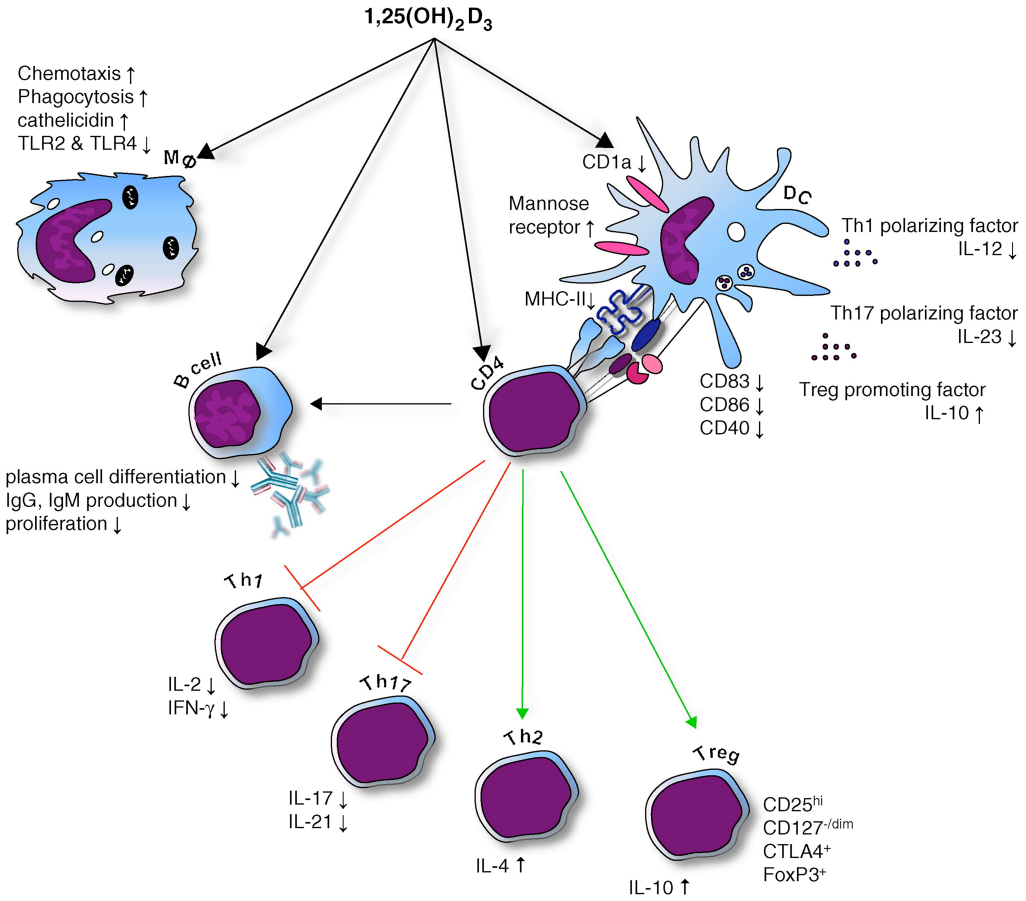



**Figure 1 Effects of vitamin D on innate and adaptive immune cells.**

Vitamin D increases antimicrobial and inflammatory responses in innate immunity while increasing regulatory functions of the adaptive arm. Figure reproduced from (Baeke et al., 2010) with modified text.

## THE VITAMIN D RECEPTOR

VDR is found in the nucleus as a homodimer in the absence of $1,25(OH)_2D_3$. The VDR protein can bind to the three isoforms of RXR and acquires the active conformation when bound to $1,25(OH)_2D_3$. Importantly, loss of function mutations in the VDR gene cause hereditary vitamin D resistant rickets type II (Kitanaka, 2010; Tiosano et al., 2013). There is a specific consensus motif to which VDR binds

although the extent and determination of are an area of active research (Nishikawa et al., 1994; Ramagopalan et al., 2010; Wang et al., 2005).

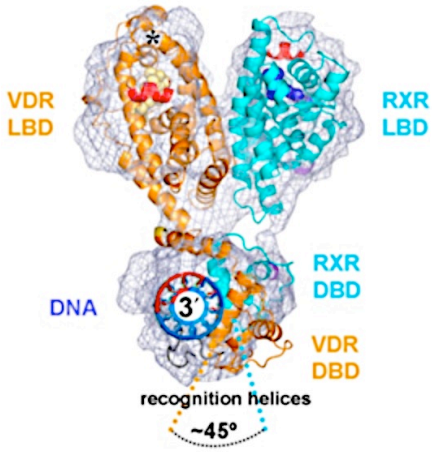



**Figure 2 Crystal structure front view model of the RXR/VDR/DR3 DNA nuclear receptor complex as seen from the 3′ -end of the DNA.**

The three nucleotide spacer (DR3) between the VDR and RXR half-sites causes the DBDs to be rotated by 45° with respect to each other. DBD = DNA binding domain, RXR = retinoid X receptor, VDR = vitamin D receptor; LBD = ligand binding domain. Magenta = cryo-EM map; blue = DNA; green = first half-site of the response element; red = second half-site; cyan = DBD of RXR; orange = DBD of VDR; yellow = $1,25(OH)_2D_3$; blue = 9-cis retinoic acid; star = helices H2, H3n and the β-sheet of VDR; red = VDR trans-activation helix H12. Helix H12 allows recruitment of chromatin-modifying co-regulator proteins. Figure adapted from (Orlov et al., 2012).

Two recent studies provide very detailed structural information using the classic consensus sites (Orlov et al., 2012; Zhang et al., 2011). VDR as a heterodimer with

RXR forms a DNA interacting complex with DNA binding domains in each protein (Figure 2). VDR/RXR's consensus binding element is the half site 5'-AGGTCA-3' with 3 spacer nucleotides and a direct repeat of the half site (so called DR3) (Orlov et al., 2012). RXR will bind the 5' half site. Several variations and spacers of more than five elements have been described (half site only, hexanucleotide everted repeats with 6n spacers) (Pike et al., 2010). The factors that determine specificity of target regulation are still incompletely understood. These can affect the interplay with co-factors and co-repressors that ultimately determine gene regulation (Figure 2 and Figure 4).

Multiple polymorphisms and mutations that lie within or affect *VDR* have been described and several disease associations have been reported. Mutations that truncate VDR and cause hereditary vitamin D deficiency rickets have been described above. Four single nucleotide polymorphisms in particular have been widely studied and have been variably associated to several diseases including MS and tuberculosis (Figure 3) (Andraos et al., 2011; Hotchkiss et al., 2013; Smolders et al., 2009).

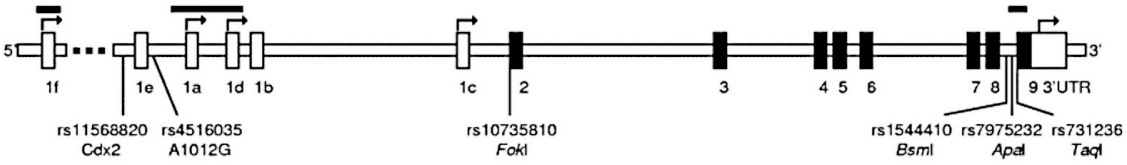



**Figure 3 The human vitamin D receptor gene showing polymorphisms associated to disease (tuberculosis and multiple sclerosis).**

The region shown corresponds to chr12:46,524,238–46,525,237 on the negative strand (UCSC release hg18, March 2006). VDR spans 105 kb in the region q13.11. Exons marked 1f to 1c are largely non-coding and are indicated by open bars. Black bars indicate coding exons 2 to 9 with transcriptional start sites marked by curved arrows. Studies independently looking at MS or tuberculosis have shown associations of susceptibility to these diseases with *ApaI* (rs7975232), *BsmI* (rs1544410), *FokI* (rs10735810) and *TaqI* (rs731236). Figure adapted from (Andraos et al., 2011).

## VDR IN EVOLUTION

The discovery of VDR in the jawless primitive fish (lamprey) led to the conclusion that the vitamin D system originated before the development of calcified structures (Whitfield et al., 2003). Ultraviolet irradiation in invertebrates, fungus and plants generates vitamin $D_2$ and thus vitamin D may have developed as a system to help prevent damage to nucleic acids and proteins (Holick, 2011). VDR's structural similarity to other nuclear receptors involved in bile acid and xenobiotic receptors raises the possibility that its initial importance related to detoxification and later evolved to calcium homeostasis and immunity (Bouillon et al., 2008). The VDR gene

is conserved in dog, cow, mouse, rat, chicken, zebrafish, fruit fly, and mosquito and it seems likely that *VDR* has been positively selected in humans (Grossman et al., 2010). It is thought that skin colour evolved to address vitamin D synthesis (Jablonski et al., 2000). Vitamin D deficiency leading to disorders such as rickets-induced pelvic fractures can have fatal consequences and may have driven selection to some extent (Neer, 1975).

## GENOMIC ACTIONS OF THE $1,25(OH)_2D_3$ –VDR/RXR COMPLEX

The VDR complex regulates transcription of many genes in a range of tissues (Heikkinen et al., 2011; Wang et al., 2005). Vitamin D exerts most of its influence through VDR and binding to DNA elements although non-genomic actions have been described (Norman, 2006; Wu et al., 2010). Upon intracellular binding of $1,25(OH)_2D$, VDR forms a heterodimer with RXR and binds to specific genomic sequences (VDREs) to regulate transcriptional responses (Pike et al., 2010). The $1,25(OH)_2D$-VDR-RXR complex recruits general transcription factors that regulate gene expression (Figure 5). Depending on the location of binding, the VDR complex may help initiate transcription at promoter sites or modulate expression at distant enhancer locations (Pike et al., 2010; Ramagopalan et al., 2010).

### *Genome-wide binding*

A joint study by the J Knight Lab and Ebers Lab determined the first genome-wide map of VDR binding using chromatin immunoprecipitation followed by massively parallel DNA sequencing (ChIP-seq). Following $1,25(OH)_2D_3$ stimulation of human lymphoblastoid cell lines, 2776 genomic locations bound by VDR were identified (Ramagopalan et al., 2010). Increased VDR binding in intronic and intergenic regions

compared with the basal state was observed as well as a correlation between the number of motifs within an interval and the strength of the interval. A motif closely resembling the DR3 and DR4 consensus motifs was found in 60% of binding locations (Figure 4).

Interestingly, comparison of VDR binding distribution with ENCODE data showed a greater enrichment of VDR binding sites at H3K4me3 (associated with active transcriptional start sites; TSS) and H3K27ac (associated with active enhancers) locations as compared with H3K4me1 (associated with cis-regulatory regions distal and proximal to the TSS). Moreover, the enrichment of VDR occupancy in DNase I–hypersensitive sites (associated with regulatory sites) and CTCF (associated with insulators) was also significant.

In relation to vitamin D–responsive genes assessed in the same study, 23% contained a VDR interval near the TSS. Excluding these, vitamin D-responsive genes had a VDR interval on average at a distance of 66.6 kb away from the TSS, in contrast to intervals in non-responsive genes found at 352.8 kb. Finally, regions of positive selection among individuals of Asian and European descent also showed significant enrichment of VDR intervals (Ramagopalan et al., 2010).

This data shows that VDR binding sites occur at specific regulatory regions within the epigenetic landscape and likely contribute to widespread regulatory functions that have played a role in evolution. Further studies of VDR binding in other cell lines have since been published in human monocytes (THP-1 cell line) (Heikkinen et al., 2011), in human colorectal cells (LS180 cell line) (Meyer et al., 2012) and in human

hepatic stellate cells (LX2 cell line) (Ding et al., 2013). These studies have shown that VDR binds to between 1,600 and 6,200 VDREs with similar motifs but a low overlap of binding sites between cell types (Carlberg et al., 2013).


**Figure 4 Sequence logo for VDR ChIP-seq binding intervals after 1, 25 $(OH)_2D_3$ stimulation of lymphoblastoid cells.**

The top-scoring match found by MEME resembles the VDR motif consensus. In chapter three I present two polymorphisms located at conserved motif positions that disrupt VDR binding. Figure reproduced from (Ramagopalan et al., 2010).

***Gene expression studies***

A variety of technological platforms have been used to study gene expression after vitamin D stimulation, mostly using $1,25(OH)_2D_3$. Studies using genome-wide methods have found that the expression of an important number of genes (0.8 to 5% according to the study and cell type) is modulated by vitamin D (Bouillon et al., 2008; Heikkinen et al., 2011; Ramagopalan et al., 2010; Wang et al., 2005). Few tissues appear not to express *VDR* or *CYP27B1*, or only express them at low levels, including mature striated muscle cells and Purkinje cells in the cerebellum (Eyles et al., 2005). In a recent study, our group observed 229 genes with significant changes in

expression in response to vitamin D using naïve B cells (Ramagopalan et al., 2010). These and several other studies revealed that vitamin D is involved in a variety of functions including immune and inflammatory responses; cellular growth, differentiation and apoptosis; DNA repair and oxidative stress; and membrane transport and cell adhesion (Ramagopalan et al., 2010; Stewart et al., 2004). These studies have also highlighted that vitamin D acts in multiple cell types including bone, kidney, intestine, immune and smooth muscle cells, among others (Hsu et al., 2006).

***Gene regulation***

Activated vitamin D (calcitriol, $1,25(OH)_2D$) exerts control over gene expression upon binding to VDR and forming the VDR/RXR complex (Figure 5). This in turn binds to genomic VDREs and recruits co-regulatory protein complexes (McKenna et al., 2002) with loss of co-repressive factors (Tagami et al., 1998). It is likely that unbound VDR is kept inactive by co-repressors (SMRT, NCoR, Alien, sin 3) that deacetylate histones (Tagami et al., 1998). VDR binding to VDREs appears to occur in cyclical waves (Kim et al., 2005) although the nature of transcription factor kinetics is not well understood (Fullwood et al., 2009). Both positive and negative gene regulation are thought to be exerted by vitamin D with poorly understood chromatin interactions occurring (Fujiki et al., 2005; Kitagawa et al., 2003; Villagra et al., 2006).

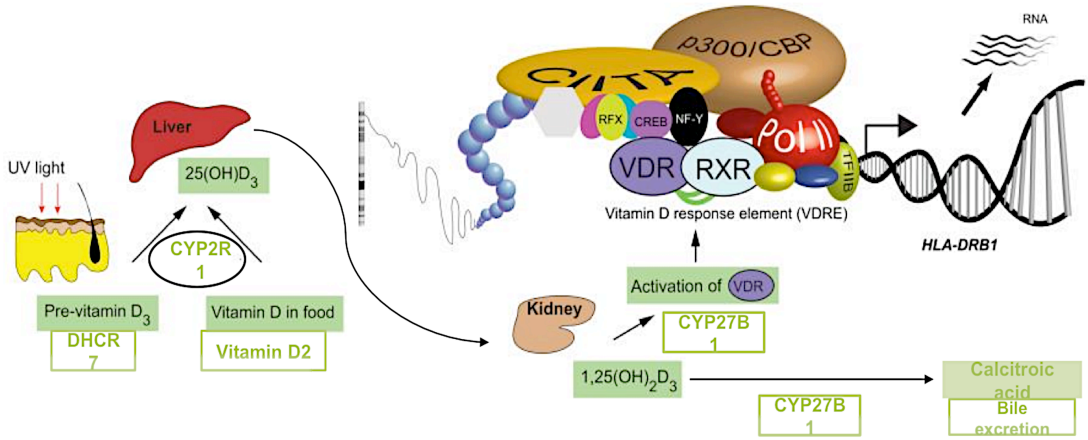



**Figure 5 Illustration depicting vitamin D metabolism and molecular function:**

7-dehydrocholesterol (DHCR7) is converted to pre-vitamin $D_3$ on ultraviolet B exposure initiating a series of enzymatic conversions that can eventually lead to active vitamin D (1, 25 $(OH)_2D_3$) regulating gene expression. Here a proximal vitamin D response element (VDRE) at the *HLA–DRB1* gene that confers vitamin D responsiveness through binding of activated vitamin D to its canonical receptor in heterodimer formation. VDR/RXR are shown as well as key vitamin D forms, enzymes and general transcription factors. Figure adapted from (Handunnetthi et al., 2010).

***Positive gene regulation***

A complex cascade of events initiates when proteins of CBP/p300, p160 and steroid receptor co-activator families are recruited to the VDR/RXR ligand bound-VDRE site (Rachez et al., 2000). These co-activators open the chromatin structure by acetylating histones (Spencer et al., 1997) which allows the DRIP/TRAP complex (VDR interacting protein/$T_3$ receptor auxiliary protein) to be recruited directly to the DNA bound VDR/RXR-ligand dimers (Vanhooke et al., 2004). Once this occurs, basal

transcription factors such as RNA polymerase II can be recruited to transcriptional start sites and specific gene expression occurs.

***Negative gene regulation***

Negative VDREs (nVDREs) closely resemble the consensus VDRE and have been found in the *PTH*, *CYP27B1* and *PTHrP* promoters. Negative VDREs can bind VDR homodimers or VDR/RXR heterodimers, may contain Ebox type motifs and may bind the VDR/RXR dimer with the opposite orientation as positive VDREs having the VDR occupy the 5' half site (Kim et al., 2005; Murayama et al., 1999). The exact mechanisms as to how nVDREs work are not known but replacement of CBP/p300 HAT co-activators with NCoR/SMRT co-repressors and histone deacetylases is thought to occur (Fujiki et al., 2005; Murayama et al., 1999). Alternatively, vitamin D may act indirectly by inhibiting upstream molecules and thus downregulating target genes as is the case with NFATp/AP-1 and NF-kappaB acting on IL-2 and IL-12, respectively (Alroy et al., 1995; D'Ambrosio et al., 1998). Interestingly, vitamin D and VDR resistance have been described in New World Primates and in a single human patient with a normal VDR. High expression of a vitamin D responsive element binding protein was found that is homologous to members of the heterogeneous nuclear ribonucleoproteins family (Chen et al., 2003). It is postulated that these proteins negatively modulate the binding of steroidal nuclear receptor transcription factors.

## Vitamin D in autoimmune and inflammatory diseases

Vitamin D deficiency has been associated to a number of diseases but causality and molecular mechanisms remain poorly understood. Studies showing that vitamin D is important in both the adaptive and innate immune arms are accumulating (Adorini et al., 2008; Hewison, 2011). Recent investigations collectively suggest that vitamin D may be causal in some autoimmune disorders, notably type 1 diabetes and MS (Badenhoop et al., 2012; Giovannoni et al., 2007). Multiple lines of evidence including human genetic association studies, epidemiological data, inadequate innate and adaptive immunity during vitamin D deficiency reversible with repletion, and animal model studies indicate this (Badenhoop et al., 2012).

Similarly, the long-standing hypothesis of an involvement of vitamin D in antimicrobial and inflammatory responses is gaining credence through molecular studies and pilot clinical trials (Martineau et al., 2011; Olliver et al., 2013). An anti-tuberculous macrophage mechanism involving interferon gamma and the antimicrobial peptide cathelicidin, which is dependent on sufficient 25(OH)D levels, has been described (Fabri et al., 2011; Martineau et al., 2007). Interferon gamma and VDR expression also appear to be relevant in the pathogenesis of MS by mediating reduction of pathogenic T cells (Spanier et al., 2012). Interferon dysregulation has been observed in lepromatous patients (Teles et al., 2013) and may constitute a balancing mechanism that regulates innate and adaptive responses that is dependent on adequate levels of vitamin D and functional vitamin D metabolism genes. Other pathways, including additional anti-bacterial responses and regulation of autophagy

(Hewison, 2011), may be equally reliant on vitamin D homeostasis and much remains to be understood. In this section I aim to highlight important advances and areas of need in our understanding of vitamin D biology in immune and inflammatory conditions by focussing on MS and sepsis as examples.

## MULTIPLE SCLEROSIS: A COMPLEX DISEASE INFLUENCED BY VITAMIN D

### *Definition and epidemiology of MS*

A complex autoimmune disorder that targets the central nervous system, MS is the most prevalent neurological disease causing disability in young individuals. MS susceptibility is influenced by genetic and environmental factors (Ascherio et al., 2007; Sadovnick et al., 1996). It is generally considered autoimmune and inflammatory, causing progressive disability (Compston et al., 2008). Incidence and prevalence increase with latitude of childhood and adolescent habitation (Kurtzke et al., 1970; Willer et al., 2005) and latitude gradients correlate inversely with ultraviolet B (UVB) exposure (Handel et al., 2010; Simpson et al., 2011).

Vitamin D and its deficiency have been associated to MS in epidemiological, animal models, clinical and molecular genetic studies. The incidence and prevalence of MS vary according to geographic location (Handel et al., 2010). Prevalence rates of 60 per 100,000 or more are found in Europe, southern Canada, northern United States, New Zealand and southeast Australia. The Orkney Islands (northern Scotland) present the highest reported rate with 300 per 100,000 with substantially lower rates in countries found nearer to the equator (Kingwell et al., 2013). Racial differences with particular genetic backgrounds may partly explain geographic variation, as susceptibility is greater in white populations while individuals with Asian, African, or American Indian origin have lower risk (Handel et al., 2010). Environmental influences in MS susceptibility and progression are strongly implicated however and gene-environment interactions are thought to occur (Giovannoni et al., 2007).

***Clinical presentation and diagnosis of MS***

MS is heterogeneous with variable clinical presentations and pathologic findings. Important disease mechanisms involve inflammation, demyelination and axon degeneration. The cause of MS remains unknown and the currently accepted theory is that self-reactive T cells mediate an inflammatory and autoimmune condition that is specific to the central nervous system (Frohman et al., 2006; Sawcer et al., 2011). MS has a gender ratio with more women than men affected with a 2-3:1 ratio. This has been increasing over the last fifty years and suggests gene-environment interactions amongst several factors (Alonso et al., 2008; Koch-Henriksen et al., 2011).

The median age of disease onset is 23 years (mean is 30 years) with incidence and prevalence varying according to population and geography (Ebers, 2008). Onset is typically poly-symptomatic over months to years. Diagnosis is often difficult and requires repeated examination as well as imaging and cerebrospinal fluid analysis for oligoclonal bands (Polman et al., 2005). No clinical findings are unique to MS although some are characteristic and common symptoms have been described (Table 17 and Table 18) (Olek, 2013). Patients are often young adults with two or more clinical episodes of central nervous system (CNS) dysfunction.

MS progression is most often of relapse and remission with acute or sub-acute onset of clinical dysfunction within days or weeks and remission over a variable period with complete or incomplete recovery (Lublin et al., 1996; Scalfari et al., 2010). Most patients (85-90%) have relapsing and remitting MS with the majority eventually entering a secondary progressive form. Patients with primary progressive MS

experience a steady decline in function from onset without acute attacks. Progressive relapsing MS from disease onset is defined by underlying decline with clinically definable relapses.

***Risk factors in susceptibility and progression of MS***

Genetic risk factors for MS susceptibility have been studied for several decades and a clear association is now evident (Ebers et al., 1996; Sawcer et al., 2011). Studies looking at MS families have observed risk varying from 3% to 23%. The Danish MS registry study used population wide data with 8205 MS patients and found that the relative lifetime risk was increased sevenfold (95% CI 5.8-8.8) among first-degree relatives (n = 19,615) (Nielsen et al., 2005). Sporadic absolute risk of MS in Danish women and men is 0.5 and 0.3% respectively while for first-degree relatives of MS patients the excess lifetime risk was 2.5% (95% CI 2.0-3.2) above. Twin studies showed that the risk of developing MS in dizygotic twins is the same as in siblings (3-5%) although the risk for monozygotic twins is between 20% and 39% (Sadovnick et al., 1993) (Figure 6).

The most consistent and important genetic region associated to MS susceptibility is the HLA-DRB1 locus of the major histocompatibility complex (MHC; also human leukocyte antigen –HLA-) (Dyment et al., 2005; Sawcer et al., 2011). Genome-wide association studies have provided increasing evidence of genetic risk at non-MHC loci of modest effect and predominantly associated to immune and inflammatory genes (Sawcer et al., 2011).

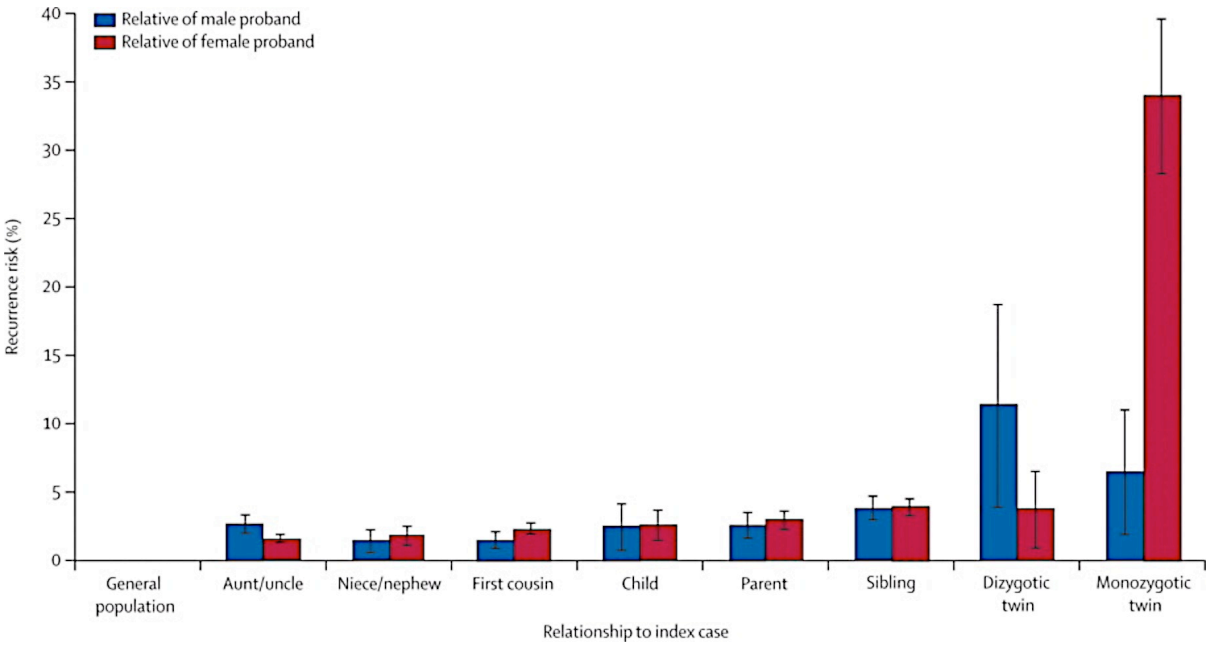



**Figure 6 Age-adjusted percentage recurrence risks for relatives of multiple sclerosis probands.**

MS risk increases with increasing shared genetic background. Data shown represent the mean with standard error bars and segregated by gender. Figure reproduced from (Ramagopalan et al., 2010) with modified text.

Many environmental risk factors have been proposed for MS including vitamin D deficiency, smoking and infectious agents. Strong evidence associating viruses directly to MS susceptibility is lacking but important observations have increasingly been made regarding Epstein Barr Virus (EBV). EBV causes infectious mononucleosis and has been postulated as a cause or trigger of MS (Disanto et al., 2013; Pender, 2009). EBV can be found in 83-90% of adults while EBV seropositivity reaches 100% among adult MS patients and is significantly different to healthy controls (Pender, 2009).

Other environmental factors such as smoking have also been suggested in MS susceptibility and progression with increased relative risks of around 1.8 for ever-smokers vs. never-smokers in Norwegian and UK populations (Hernan et al., 2005; Riise et al., 2003). Stronger evidence is required to ascertain its role however.

***Epidemiological evidence associating vitamin D to MS pathology***

The incidence, prevalence and mortality of MS increase with increases in latitude, both in the northern and southern hemispheres (Alonso et al., 2008; Handel et al., 2010; Hernan et al., 1999; Llorca et al., 2005; McDowell et al., 2010; McLeod et al., 1994). This is thought to be mediated by both genetic and environmental factors (Ascherio et al., 2007; Sadovnick et al., 1996) and evidence suggests that for the environment it is UVB radiation (UVBR) which largely mediates the latitude gradient. UVBR in turn determines vitamin D levels (Holick, 2007). Pockets of populations in northern latitudes with a lower prevalence of MS as compared to surrounding areas appear to be explained by high dietary consumption of fish, rich in vitamin D (Kampman et al., 2008); increased summer outdoor activity has also been shown to reduce risk of MS (Kampman et al., 2007).

Other lines of evidence pointing towards UVBR modifying risk have used migration studies. These have demonstrated that risk for MS changes and migrants acquire the risk of their new area of settlement (Oren et al., 2010) but only according to age of migration (Kurtzke et al., 1970). For migrants moving from areas of high to medium-low prevalence, the risk of disease was retained if they were over 15 years of age but reduced otherwise. On the contrary, migrants from low to high prevalence areas

increase their risk if they move before adolescence but not at later points in life (Kurtzke et al., 1998). Risk for MS is increased in people born in late spring as compared to those born in late autumn (Ebers, 2010; Templer et al., 1992; Willer et al., 2005). The mechanism behind this phenomenon is not clear but it is hypothesized that seasonal variations in UVBR, microbial burden or other factors mediate it. It is interesting to note however that the month of birth effect changes according to latitude (Staples et al., 2010) and is also present in many other diseases, particularly T1D which is now strongly associated to vitamin D deficiency (Kahn et al., 2009).

Munger and colleagues have provided some of the best evidence for the association of vitamin D to MS by showing in a prospective, nested case-control study among 7 million US military personnel for whom serum samples had been stored prior to disease that the relative risk for MS decreases as serum vitamin D levels increase (Munger et al., 2006). Gene-environment interactions were implicated by observed increases in monozygotic twin concordance with latitude and stronger month of birth effects when MS is familial (Islam et al., 2006). These studies have provided circumstantial evidence to further investigate how vitamin D may influence disease at a molecular level (Figure 8).

***Genetic variants associated to MS and vitamin D***

Epidemiological and clinical studies investigating latitude gradients, effects of migration, month of birth and pre-disease blood levels of vitamin D have led to more detailed genetic and molecular studies (Berlanga-Taylor et al., 2011). Genetic variants in the VDR gene in MS cases have been studied in patients vs. controls and family

based transmission test designs. Reports of positive associations in *ApaI* have been made in a Japanese population and were reproduced and extended in an Australian population (Smolders et al., 2009). VDR gene polymorphisms findings have not all been replicated in different populations however. *ApaI* (rs7975232) and *TaqI* (rs731236) SNPs associated to MS have been reported in Caucasians but have not been replicated in the same population (Smolders et al., 2009). The SNP *FokI* (rs10735810) has been associated with serum levels of 25-hydroxyvitamin D in a twin population by the Ebers lab (Orton et al., 2008) and was later confirmed in a study using patient cases vs. healthy controls by another group (Smolders et al., 2009). Similarly, a polymorphism in the gene *CYP27B1* was found in a Caucasian population (rs703842) (ANZgene, 2009) and later confirmed by a different group (Sundqvist et al., 2010).

The most recent GWAS of susceptibility to MS showed that common SNPs near *CYP24A1* and *CYP27B1* increase risk (Sawcer et al., 2011). Further, a recent study has shown that individuals heterozygous for rare *CYP27B1* mutations carry increased risk for MS while homozygous individuals develop vitamin D dependent rickets (Ramagopalan et al., 2011). Finally, the main MS risk locus, HLA-DRB1*15 has a conserved VDRE that appears non-functional and variant in non-associated HLA-DRB1 alleles (Ramagopalan et al., 2009).

***Interactions in MS: vitamin D, EBV and interferon treatment***

Several interactions have been shown between MS risk factors. Importantly, an EBV protein, EBNA-3, has been reported to bind to VDR and block vitamin D responsive

genes inhibiting apoptosis and growth arrest (Yenamandra et al., 2010). A recent study showed that low vitamin D levels and high anti-EBNA-1 antibodies are present at least two to three years before disease manifestation in MS patients (Decard et al., 2012). These were not correlated but showed significant differences when compared to matched healthy controls. VDR and oestrogen interact at various levels and understanding these may help elucidate the gender ratio and increasing prevalence of MS in women. A direct interaction between vitamin D and smoking is not clear and dissecting lifestyle and other confounding factors remains challenging. Interestingly, vitamin D is known to regulate interferon gamma receptor 2 expression and interferon gamma is necessary for brain expression of *Vdr* in mice (Spanier et al., 2012).

Further studies are required to fully understand the implications of these observations in terms of disease mechanisms and patient response to treatment. Other interactions, such as epistasis or gene-hormone events, are only beginning to be investigated.

***A molecular basis for vitamin D in MS: genes, pathways and cell types***

Several recent studies have provided a molecular basis for the epidemiological and clinical observations of the relationship between vitamin D and multiple sclerosis. Although many questions remain unanswered, it now seems clear that vitamin D, through VDR, interacts at a genome-wide level with genetic risk variants for MS, including HLA-DRB1*15:01 and other HLA-DRB1 alleles (Ramagopalan et al., 2011; Ramagopalan et al., 2010; Ramagopalan et al., 2009). Additionally, many genes near MS genetic risk variants have VDR binding sites within or nearby more than expected by chance, including non-MHC genes such as *CD40* and *IRF8*.

Vitamin D exerts a wide range of effects and regulates hundreds of genes including pathways related to membrane transport and cell adhesion; DNA repair and oxidative stress; cellular growth, differentiation and apoptosis; and immune and inflammatory responses. Several molecules are known to be regulated by vitamin D and could play an important role in MS such as NFkB, IL-10 and IL-17. These molecules have important functions in immune and inflammatory responses and play a role in vitamin D responses in the MS animal model experimental autoimmune encephalomyelitis (Carlson et al., 2013).

The effect of vitamin D on MS GWAS implicated genes has not been systematically explored however. The main risk locus, *HLA-DRB1*1501* appears to be directly regulated by calcitriol (Ramagopalan et al., 2009). Other genes implicated in MS physiopathology or susceptibility such as *TOB1, TSFM* and *TNF* family members may be modulated by vitamin D and could help explain possible effects of vitamin D supplementation on MS progression.

Vitamin D may act at various points in development and on different tissues. Identifying the existence and understanding the role of trans-generational effects of vitamin D deficiency is a key question that highlights the complexity of disease and the importance of understanding risk factors across multiple dimensions. Determining what genes, pathways and cells are important in MS development, onset, course and severity requires much further study.

***Disease mechanisms in MS: an immunological role in causation and progression***

Several studies have provided an autoimmune basis for MS although a direct causal link has not been shown. An increasing overlap of genetic susceptibility loci to both MS and type 1 diabetes as discovered by GWAS is becoming apparent and lends support to the theory of shared genetic risk for autoimmune conditions (Consortium, 2009). T and B cells as well as other immune cell types are thought to be important in MS pathophysiology and are the target of several forms of treatment (Compston et al., 2008). The main genetic risk locus, HLA class II, has key immunological functions and a large recent GWAS of MS susceptibility strongly implicates genes involved in T-helper cell differentiation pathways (Sawcer et al., 2011) (Figure 7).

Abnormal immune peripheral responses in MS patients have been observed including the failure of regulatory lymphocytes to suppress effector cells (Viglietta et al., 2004). Different inflammatory and immune mediated mechanisms have been postulated to be directly involved in the pathogenesis of CNS specific injury in MS: CNS T-cell and macrophage infiltration with ensuing local inflammation and surrounding tissue damage; antibody and complement-mediated immune responses against oligodendrocyte cell lineage and myelin; vascular damage and/or mitochondrial dysfunction secondary to inflammation and macrophage toxins, respectively, leading to hypoxia-like injury; and genetic variants predisposing oligodendrocytes to primary immune targeting (Compston et al., 2008). The formation of the sclerotic plaque in the CNS is the pathognomonic endpoint of demyelinating disease involving a complex process of inflammation, demyelination and remyelination, oligodendrocyte

depletion and astrocytosis, and neuronal and axon degeneration (Compston et al., 2008).

Specific antigens triggering the immune response in MS remain equivocal although myelin proteins, αB crystalline and neurofascin have been implicated. Susceptibility and progression in MS are thought to be the consequence of cumulative axon loss initiated by self-reacting immune cells and maintained by inflammatory responses in genetically susceptible individuals with complex lifetime environmental exposures (Compston et al., 2008; Handel et al., 2010) (Figure 8).

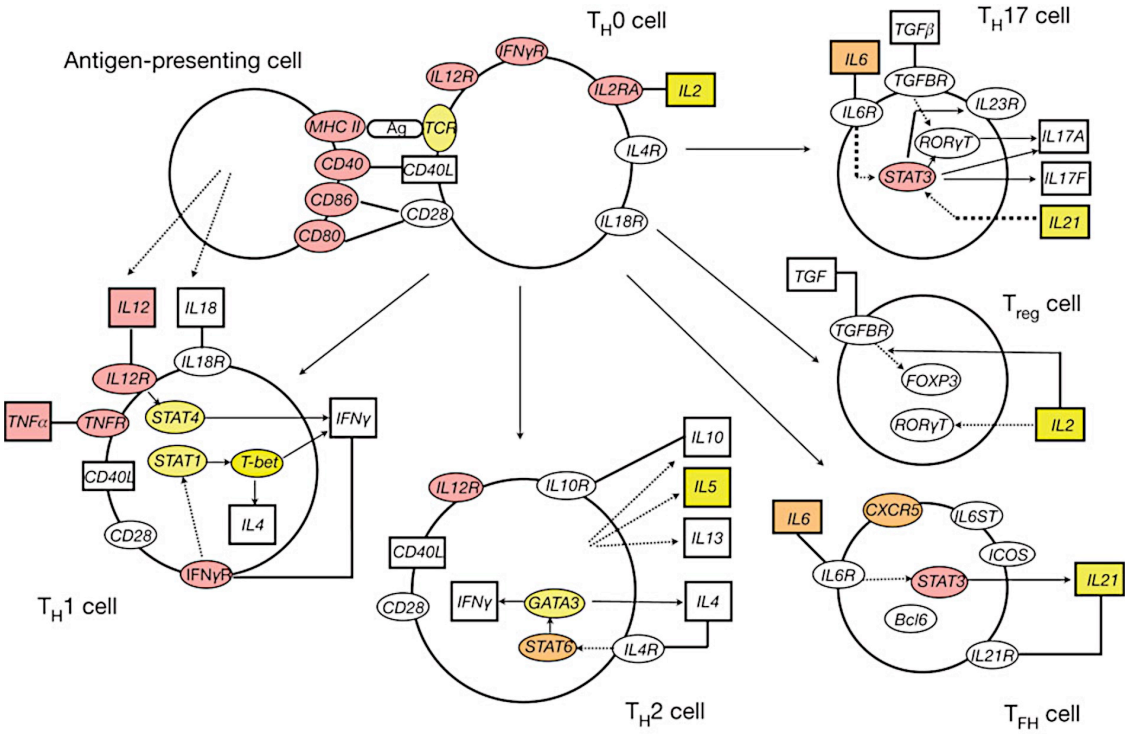



**Figure 7 The T-helper cell differentiation pathway showing genes close to GWAS identified variants for MS susceptibility.**

The authors used Ingenuity Pathway analysis to identify known pathways involving genes near genetic variants associated to MS. Nodes are coloured

according to the strength of association: red, in bold or grey strongly associated with genome-wide significance; orange = discovery P value < 1 × 10–4.5 and consistent replication; yellow = discovery P value < 1 × 10–3. Figure reproduced from (Sawcer et al., 2011) with modified text.

***Treatment and prognosis of MS***

The natural history of disease in relapsing remitting MS patients has been dissected in a large, geographically based, systematically ascertained cohort of ~28,000 patient-years of untreated multiple sclerosis from the London Ontario Database across more than 10 studies (Degenhardt et al., 2009; Weinshenker et al., 1987). These analyses allowed a detailed examination of prognostic factors useful for clinical trial design and interpretation (Degenhardt et al., 2009). Early relapse rate with other negative prognostic indicators may be useful outcome surrogates for therapeutic and preventative trials (Scalfari et al., 2010). These factors could be the best measures of therapeutic relevance although carrying out long term clinical trials for these end-points is costly and most studies use relapse rates at various time-points.

In recent years a number of drugs have become available that appear to modify disease course. These have mainly been targeted for relapsing-remitting MS and include interferon beta drugs and glatiramer acetate as first line agents (Goodin et al., 2012). The mechanism of action of interferon beta-1b in MS is unknown but it is thought to modulate immune responsiveness. It is one of the oldest drugs in the market and recently has been shown to confer long-term survival in MS patients

(Goodin et al., 2012). Glatiramer acetate is antigenically similar to myelin basic protein and composed of random polymers of four amino acids. It has two mechanisms of action proposed: binding to MHC molecules and competition with myelin antigens, and induction of T helper 2 suppressor cells that cause bystander suppression in the brain with anti-inflammatory cytokine secretion profile (Arnon et al., 2004).

Most drugs currently in use and in development target the immune system and important efforts are under way to test novel agents as well as combinations and supplementation with vitamin D (Carlson et al., 2013; Naismith et al., 2010; Olek, 2013).

***Vitamin D as a marker of disease in MS***

Few studies have assessed whether vitamin D serum levels can be used to help predict disease risk or progression in MS. The prediction of conversion to MS from radiologically or clinically isolated syndromes (RIS or CIS respectively) may be possible if serum vitamin D levels are combined with other data e.g. anti-EBNA-1 titres and genetic variants.

***Vitamin D as prevention and therapy for MS***

Trials of vitamin D in MS have been completed and others are on-going (Carlson et al., 2013). These have so far been small and likely underpowered. Trials with large sample sizes and carefully designed clinical outcomes are necessary to test the effectiveness of vitamin D as a therapeutic agent in MS.

Recently concluded studies evaluating high vs. low levels of vitamin D in blood have shown differences when comparing relapse rates, new MRI detected brain lesions and disability scores. Of the recently completed studies, one by Mowry and colleagues had the largest sample size (n=469) and longest duration (5 years) (Mowry et al., 2012). This, along with the several studies of vitamin D levels in blood that preceded it, lend further support to help justify clinical trials and inform design.

Clinical trials testing vitamin D supplementation outside of MS have been largely disappointing so far, particularly for infectious respiratory diseases, although vitamin D as adjunct therapy in tuberculosis appears promising (Martineau et al., 2011). Negative trials have raised the issue that findings gleaned from observational epidemiological studies may have arisen from confounding. Recent reports have highlighted the general lack of evidence stemming from randomised clinical trials and prospective studies in vitamin D research (Ross et al., 2011). Adequately designed and powered studies are necessary to address the benefit of supplementing with vitamin D, including studies in MS patients and at-risk individuals. Study designs for vitamin trials are particularly challenging. Baseline levels, achieved serum levels, population to treat (CIS, RIS or MS patients for instance), duration of treatment, dose, formulation (1, 25 $(OH)_2D$ vs. 25 (OH)D; $D_2$ vs. $D_3$), amongst several other factors, require careful consideration (Berlanga-Taylor et al., 2013).

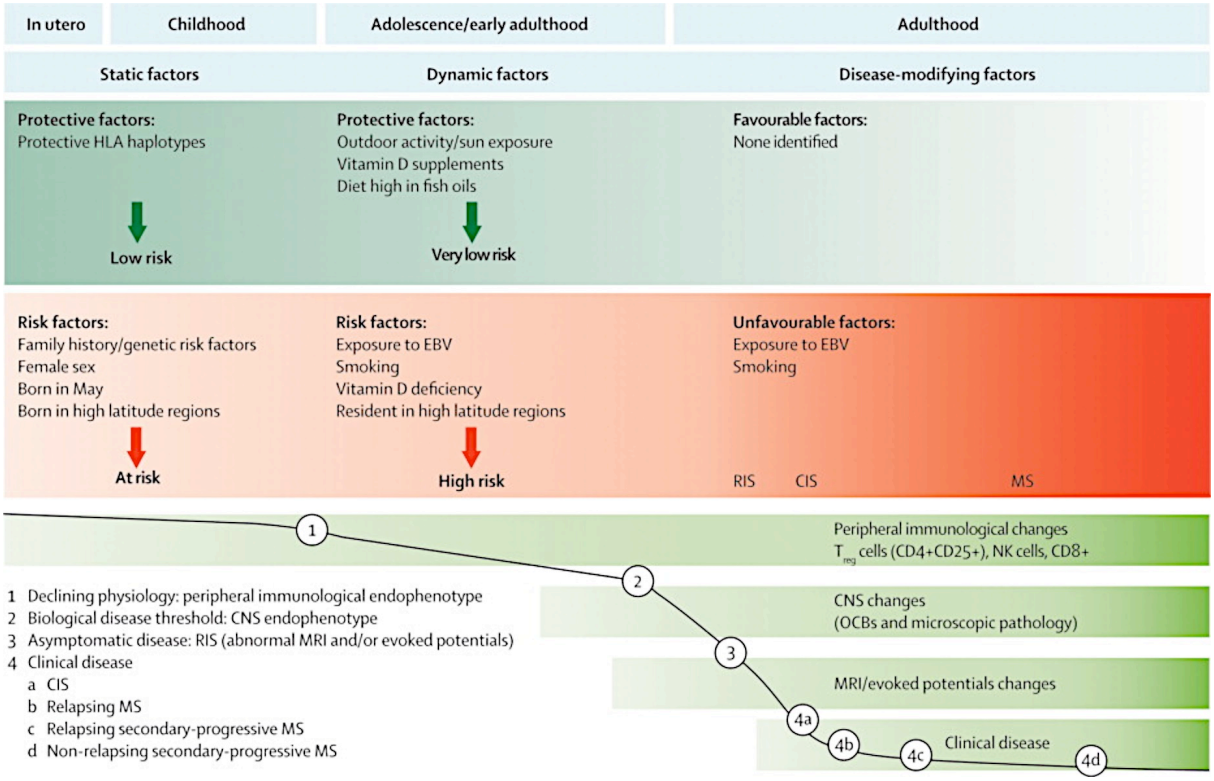



**Figure 8 Suggested causal pathway for multiple sclerosis integrating spatio-temporal effects from perinatal to adult development.**

Genetic, trans-generational, month of birth and environmental effects are likely to act at diverse moments (Handel et al., 2010). These may continue, are reinforced or exacerbated by secondary exposures during childhood and adolescence. Migration and environmental exposure to viruses, smoking and vitamin D may play a variable role in protection and susceptibility over the same period. Immunological changes may become apparent sub-clinically (1) and over time lead to overt immunologically driven changes in the CNS identified by imaging and other tests (2 and 3). Demyelination ensues following this and clinical manifestations become apparent as disease-modifying factors interact and disease progresses (4). CIS=clinically isolated

syndrome; EBV=Epstein-Barr virus; MS=multiple sclerosis; NK cells=natural killer cells; OCB=oligoclonal bands; RIS=radiologically isolated syndrome. Trig cells=T-regulatory cells. Figure reproduced from (Ramagopalan et al., 2010) with modified text.

***Vitamin D and multiple sclerosis: summary and future directions***

Multiple sclerosis is a common and complex disorder with genetic and environmental factors strongly implicated in susceptibility and progression. Several important spatio-temporal points may be relevant in developing MS with gene-environment interactions suggested at key stages of disease development (Figure 8). Of the environmental factors, vitamin D has emerged as one that may act at some or all of these points and strong evidence now implicates it in MS causality. Randomised clinical trials are now underway to assess the use of vitamin D as a preventive or therapeutic agent. Many challenges remain however, including the relevant timing of treatment, trial end-point definitions and design. The timing of pathological mechanisms is poorly understood and normalising vitamin D levels post disease onset may not have measurable benefit. Regardless, evidence of vitamin D as a causal risk factor in MS is growing. MS represents a paradigm in gene-environment interactions in disease and can serve to elucidate the cellular and molecular basis behind the immune and inflammatory effects of vitamin D. Further studies to identify specific genes, pathways, targets and mechanisms behind transcriptional and cellular regulation by vitamin D are needed. In chapters three and four I show how functional genomics methods can be used to contribute to this understanding by probing the impact of genetic variation on VDR binding and the relationship of genotype, gene

expression and blood vitamin D. I provide insight that is relevant for MS and other autoimmune disorders with novel observations at candidate loci and through genome-wide analyses. I then apply some of these observations and approaches in sepsis, a complex inflammatory disorder in which vitamin D deficiency has rarely been studied previously.

## SEPSIS SYNDROME

### *Definition and diagnosis of sepsis*

Host response to infectious agents is a highly dynamic and complex process. Its objective is to locate and control invading microbes, limit their spread and eliminate or contain them. Repair is fundamental to this as the inflammatory response causes damage to surrounding tissues. If microbial spread or inflammatory damage is not stopped it can cause systemic effects that affect remote and otherwise healthy tissue leading to sepsis and possibly to multiple organ system dysfunction (MODS). Sepsis is thus defined as a dysregulated, systemic, non-resolving inflammatory response to infection with complex dynamic pathophysiological mechanisms in which there is a continuum from sepsis to septic shock and MODS (appendix Table 19) (Quesada et al., 1995; Vincent et al., 2013).

Sepsis clinically presents as a spectrum from systemic inflammatory response syndrome (SIRS) to refractory septic shock (Ravid et al., 1993). Its pathophysiology is complex and highly dynamic with the release of pro- and anti-inflammatory mediators, interplay between innate and adaptive immune cells and damage to tissue remote to the initial site of injury or infection (Bosmann et al., 2013; Quesada et al., 1995). Sepsis has a high mortality and currently therapeutic options are limited (Cohen et al., 2012). It was initially defined by a consensus panel and was updated in 2001 and again in 2012 (http://www.survivingsepsis.org) (Dellinger et al., 2013; Levy et al., 2003). Diagnosis is based on the presence, documented or suspected, of an infection and systemic manifestations that include general, inflammatory, tissue

perfusion, organ dysfunction and haemodynamic variables to consider (appendix Table 19) (Neviere, 2013).

***Epidemiology of sepsis***

The incidence of sepsis has increased dramatically over the past four decades. Estimates indicate that it has incremented 10-fold from the 1970's in the USA (Martin et al., 2003) and Europe (Harrison et al., 2006). This is thought to be due to increasing elderly populations, immunosuppression and multidrug-resistant infection. Sepsis varies by ethnic group, season and geographic location. Incidence and case fatality rates for sepsis are significantly higher during the winter season in the USA and show a greater seasonal variation in the Northeast region (Danai et al., 2007). Danai and colleagues (2007) observed variations in seasonal respiratory infections which paralleled seasonal differences but did not fully account for changes suggesting other environmental influences may be important.

Ethnic differences have also been shown with a higher incidence in non-white populations and higher mortality rates in black men (Martin et al., 2003). The mechanisms behind these observations are not clear and many factors, including socio-economic, genetic and environmental may be at play. Disease severity also appears to be increasing with a higher proportion of patients presenting at least one dysfunctional organ (Rangel-Frausto et al., 1995).

Although sepsis mortality rates appear to have decreased over the last decade they remain high, ranging between 20 to 50 per cent in the western world (Martin et al., 2003; Padkin et al., 2003). Mortality increases along the SIRS/sepsis spectrum with

reported rates of 7, 16, 20, and 46 per cent for SIRS, sepsis, severe sepsis, and septic shock respectively (Rangel-Frausto et al., 1995). Long term outcomes following an episode of sepsis are compromised and although most deaths occur within the first six months, mortality remains elevated at one-year (Perl et al., 1995; Sasse et al., 1995) with longer term survivors suffering a persistent decrement in quality of life (Nesseler et al., 2013; Winters et al., 2010).

***Susceptibility and prognosis of sepsis***

Gram positive and negative bacteria remain the most frequently identified in patients with sepsis with increasing numbers of cases involving fungal and drug resistant infections (Dellinger et al., 2013; Martin et al., 2003). Nosocomial infection, bacteraemia, increasing age (>64 years), immunosuppression, diabetes and cancer, and community acquired pneumonia are known risk factors for sepsis susceptibility.

Genetic factors are also thought to play a role in susceptibility and outcome with epidemiological studies in families, monogenic disorders and candidate variant studies showing increase in the risk of infection (Netea et al., 2011; Sands et al., 1997). GWAS of susceptibility and outcome in severe sepsis and septic shock are currently being carried out in large European cohorts (GenOSept, https://www.genosept.eu/index.ecrf).

Genetic association to poor outcome has been studied using candidate SNP approaches testing many biologically plausible candidates. Replication has been inconsistent however and a recent GWAS for survival outcome of ~3700 European

patients with severe sepsis or septic shock was not able to replicate any previous findings (Dr. Anna Rautanen, University of Oxford, personal communication).

Many studies have looked at sepsis outcome with few predictors being consistent. Microbial components and toxins, excess pro-inflammatory cytokines, complement activation and host characteristics likely play a role in determining outcome from SIRS and sepsis. Increasing age has been clearly associated to poorer outcomes in multiple studies and likely reflects its relationship to comorbidity, immunological status, malnutrition, use of nursing homes, medical interventions and exposure to nosocomial pathogens (Neviere, 2013). The host inflammatory response, as measured by hypothermia (body temperature <35.5C) or lack of fever, and leukopenia (white cell count < 4000/$mm^3$), appears to play a determining role (Peres Bota et al., 2004). Non-urinary site of infection, blood culture positivity, nosocomial pathogen infection, inadequate or late antibiotic therapy and inadequate or late restoration of perfusion have been associated with poorer outcomes (Brun-Buisson et al., 1996; Dellinger et al., 2013; Johnson et al., 2011; Nesseler et al., 2013; Neviere, 2013; Schmidt et al., 2013; Zahar et al., 2011). When severe sepsis or septic shock develops however, the outcome is similar regardless of the causative organism, multidrug resistance of the causative organism, infection site, or presence of bacteraemia (Zahar et al., 2011).

***Epidemiological studies of vitamin D levels in intensive care and hospitalised patients***

Recent research has highlighted the role of vitamin D as an important modulator of immunity with both anti-inflammatory and anti-microbial effects (Baeke et al., 2010; Edfeldt et al., 2010; Fabri et al., 2011; Hewison, 2011). Deficiency of vitamin D is

common and has been linked to many diseases at the epidemiological and molecular levels (Holick, 2007; Ramagopalan et al., 2010). Many hospitalised patients are vitamin D deficient and recent studies have revealed that this is severe in intensive care unit (ICU) patients (Lee et al., 2009; Thomas et al., 1998). A small study with 24 patients suggested that this might also be true in sepsis (Jeng et al., 2009).

A recent study highlighted seasonal variation in sepsis incidence and mortality, with increased rates in winter that were not fully accounted for by changes in viral respiratory infections and also suggested a possible latitude effect (Danai et al., 2007). Two previous studies investigating vitamin D levels in patients with sepsis found deficiency but were not able to detect an association to mortality (Cecchi et al., 2011; Jeng et al., 2009). These studies had small sample sizes, lacked molecular characterisation and long term follow-up however. Another study found an association between vitamin D insufficiency in patients with sepsis in the emergency department and increased sepsis severity (Ginde et al., 2011).

Four recent studies using large cohorts and computerised registries of mixed diagnostic cases found that vitamin D deficiency and insufficiency before or around the time of admission to hospital or the ICU were significant predictors of short and long term mortality and increased risk of blood culture positivity (Braun et al., 2011; Braun et al., 2011; Lange et al., 2013; Quraishi et al., 2013). These results increase the strength of the current evidence but still do not allow full assessment of the directionality of the effect due to potential confounding factors deriving from differences in the study populations compared. Moreover, these investigations were

all performed in the Boston area and record duplication bias was not clearly explained. Replication in independent cohorts, ideally in prospective studies, is therefore required.

***Sepsis pathophysiology and current paradigms***

Host innate immune cells that recognize and bind microbial components typically lead the response to an infection. Phagocytes, circulating and tissue-specific, are first activated along with the expression of both pro- and anti-inflammatory molecules that orchestrate and regulate a multitude of molecular and cellular processes. Innate immune cells recognize and bind to microbial components such as lipopolysaccharide (LPS; endotoxin) or lipoteichoic acid derived from the cell walls of Gram negative or positive bacteria, respectively (Cohen, 2002).

These components, termed pathogen-associated molecular patterns (PAMPs), are recognised by pattern recognition receptors and molecules (PRRs and PRMs) such as toll like receptor 4 (TLR4 or CD14-complex) that are cell associated or present as extracellular soluble molecules (Movat et al., 1987; Neviere, 2013). TLRs, nucleotide-oligomerization domain (NOD) leucine-rich repeat proteins, and retinoic-acid-inducible gene I (RIG-I)-like helicases are the three PRR families responsible for recognition of danger signals in innate immune cells (Cook et al., 2004).

Other host receptors are also capable of recognising foreign components as danger signals such as TREM1, MDL1, C-type lectin receptors and scavenger receptors such as CD36 (Bouchon et al., 2001; Neviere, 2013). These receptors, when bound to microbial components, trigger a cascade of events that lead to phagocytosis of

microbes and debris and to killing of microbes. Production and release of pro-inflammatory cytokines by the innate arm leads to recruitment of polymorphonuclear cells (PMNs) and additional inflammatory cells (Phillipson et al., 2011).

The response to invading pathogens is tightly regulated although poorly understood in the context of sepsis (Bosmann et al., 2013). Pro-inflammatory cytokines such as tumour necrosis factor-alpha (TNF) and interleukin-1 (IL-1) as well as other cytokines (IL-2, 6, 8, 10; platelet activating factor, interferons and eicosanoids) typically act in an autocrine fashion and increase the levels of other mediators. This results in the recruitment of more PMNs and macrophages to the site of infection (Phillipson et al., 2011).

An anti-inflammatory balance counters the production of TNF and IL-1 and suppresses the initial response by inhibiting cytokine production. IL-10, IL-2 and TGF-beta are well known regulatory cytokines that enhance self-tolerance and help eradicate invading microorganisms while reducing collateral tissue damage (Banchereau et al., 2012). Many cytokines seemingly have dual functions that are dictated by the complex environment, invading pathogen(s), timing and host characteristics (Khoo et al., 2011).

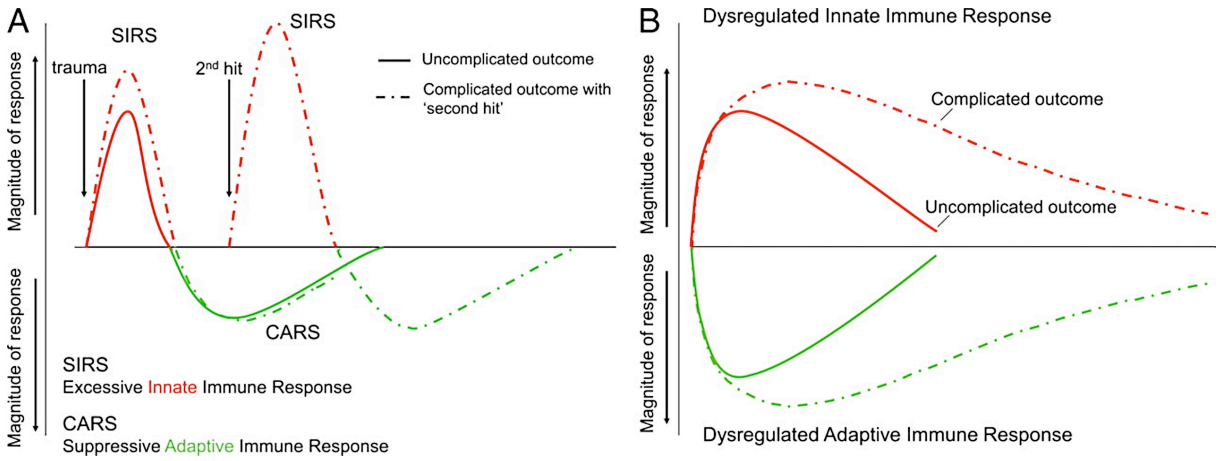



**Figure 9 (A) Current paradigm and (B) a new model in SIRS as proposed by other researchers.**

(A) Complications of trauma are thought to derive from excessive pro-inflammatory responses followed by a compensatory anti-inflammatory response (CARS) with adaptive immunity suppression. A second-hit amplifies the process and leads to recurrent SIRS and MODS. (B) Xiao and colleagues (2011) proposed that simultaneous and rapid induction of both pro- and anti-inflammatory genes occurs with suppression of the adaptive arm. Complicated vs. uncomplicated recoveries follow the same direction and differ only in magnitude and recovery. Figure reproduced from (Xiao et al., 2011) with modified text. See Figure 10 for additional information and discussion on proposed models in sepsis and areas of possible therapeutic intervention.

It is currently thought that the balance of pro- and anti-inflammatory mediators is crucial in determining outcome (Figure 9). A balance that clears the initial infectious stimuli will result in a return to homeostasis (Bone, 1996). The expected end result is

thus tissue repair and healing with clearing of invading pathogens without systemic effects. Sepsis occurs when the response to an infection exceeds the local milieu and provokes a generalized response (Bosmann et al., 2013). The factors that tip the balance are complex, dynamic and multifaceted and remain poorly understood.

A common pathway and outcome appears to occur in patients regardless of an initial infectious or non-infectious insult (e.g. pancreatitis, severe physical trauma) that leads to a systemic dysregulated response. The former are categorised as sepsis and the latter as systemic inflammatory response syndrome (SIRS) (Neviere, 2013). Current knowledge deriving from clinical and transcriptomic data indicate that this response is common in SIRS, and by extension to sepsis; is largely universal amongst human individuals and involves significant expression changes in over 80% of genes (Figure 9) (Xiao et al., 2011).

Once a generalised response is triggered, parenchymal and endothelial cellular damage ensues and underlies organ dysfunction. The mechanisms of injury are not well characterised but tissue ischemia, cytopathic injury, apoptosis and immunosuppression are thought to play important roles and interact in vicious cycles (Table 20).

Specific organ damage is common in severe sepsis and septic shock and contributes to increased mortality. Mechanisms are poorly understood and intense research is currently being undertaken in this area (Table 21) (Brown et al., 2006; Singer et al., 2004; Zhang et al., 2010). Major systems are typically affected and further contribute

to the individual's overall damage, vascular changes are often the most acute, conspicuous and threatening, and may lead to or worsen injury in other tissues.

***Treatment and recent clinical trials in sepsis***

Recent randomised clinical trials have been disappointing and currently there is no targeted therapy for sepsis (Cohen et al., 2012). Improved quality of management, faster treatment with goal directed therapy and overall ICU support improvement are likely responsible for the decrease in mortality rates over the last decade (Dellinger et al., 2013; Levy et al., 2010). The Surviving Sepsis Campaign (SSC) guidelines have been shown to be associated to a reduction in mortality following severe sepsis and septic shock (Dellinger et al., 2013; Levy et al., 2010). The SSC has generated evidence based and expert recommendations (Levy et al., 2010). Initial treatment must be directed at stabilising the patient followed by rapid identification and treatment of the site of infection (Dellinger et al., 2013).

Effective therapy targeted at controlling the immune and inflammatory response remains elusive. A recent systematic review of randomised clinical trials in adults suffering from severe sepsis or septic shock and powered for survival analysis analysed 27 studies undertaken between 1992 and 2007 (Annane, 2009). No trial moved beyond phase III except for drotrecogin alfa, a human recombinant activated protein C, which was subsequently withdrawn by the manufacturer as it failed to confirm the first positive efficacy study (Ranieri et al., 2012). Until its withdrawal in 2012, drotrecogin alfa had been the only drug licensed specifically for the treatment of sepsis. Two more recent phase II and III clinical trials separately testing a TLR-4

blocker, eritoran, and a human recombinant lactoferrin, talactoferrin, were also disappointing. The eritoran trial did not achieve the primary efficacy endpoint (Opal et al., 2013) and although talactoferrin showed promise in safety and in mortality reduction during a phase II trial (Guntupalli et al., 2013) a follow-up phase II/III was stopped because of toxic effects (McCulloh et al., 2013).

As such, there is no current targeted therapy for sepsis and options appear limited. Trials so far seem to have been confounded by sources of heterogeneity, a lack of better definitions and sub-classification of patients, inadequate estimations of the basal risk of death of enrolled patients and difficulty in controlling co-treatment (Annane, 2009).

Animal models poorly represent the human response to acute inflammation, questioning previous attempts at translation and limiting options for future work (Seok et al., 2013). Recent calls have been made to modify the current framework in sepsis research (Cohen et al., 2012). An increased effort to translate observations made in humans and identify molecular pathways and targets may allow recognition of more homogeneous subsets of patients and aid in designing effective interventions (Figure 9 and Figure 10).

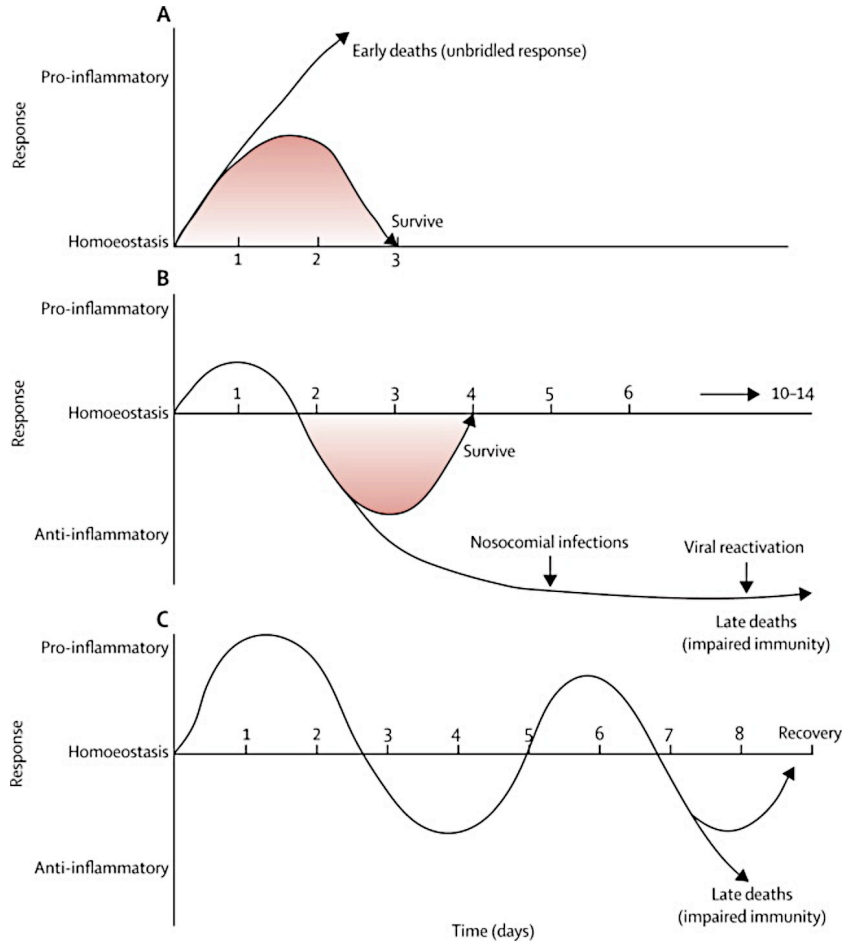



**Figure 10 Inflammatory response models in sepsis: patients may show different inflammatory responses.**

Three theoretical models are presented defining inflammatory response and possible therapeutic areas of focus (Hotchkiss et al., 2013). Otherwise healthy individuals who develop sepsis show pro- and anti-inflammatory responses that begin shortly after insult. (A) An overwhelming hyper-inflammatory reaction characterises the initial phase with early deaths commonly due to cardiovascular or metabolic derangements and rapidly developing MODS. No anti-inflammatory therapy has shown clear survival benefit although in future adequately identified patients may benefit from

currently failed drugs (Giamarellos-Bourboulis, 2013). (B) A typical patient with sepsis is an elderly individual with comorbidities that are thought to impair immune responses and thus show a reduced or absent hyper-inflammatory phase. Boosting immunity through immunoadjuvants may constitute a novel therapeutic approach in these patients. Interferon gamma therapy may act as such and is currently being tested in a phase 3 clinical trial (NCT01649921). (C) Although current data is contradictory, theoretical cycling between hypo- and hyper- inflammatory responses in a subset of patients may occur. 'Second hits' may throw patients back into a repeat hyper-inflammatory response who then recover or initiate a new cycle with deaths at any point. Biomarkers would be required in order to identify these patients and initiate therapy at the appropriate time. See also (Figure 9) above. Figure reproduced from (Hotchkiss et al., 2013) with modified text and additional references.

***Sepsis syndrome: summary and future directions***

Despite much recent progress in critical care, sepsis remains a disease of unmet need. The lack of accurate patient classification has become increasingly evident as a possible reason for failing clinical trials. Mortality remains high and current models of inflammatory processes do not fully explain disease mechanisms in the sepsis and SIRS spectrum. Many efforts have been carried out to determine risk factors in susceptibility and outcome but few have shown consistent results. This may be due to underlying heterogeneity that is not evident at the clinical level.

Large cohorts of carefully defined individuals with deep clinical and molecular phenotyping may aid patient and risk factor identification. This could then allow adequate design of therapeutic trials. Approaches using functional genomics methods tied to detailed clinical descriptions may be appropriate to answer these questions. Observational studies using these techniques could also identify, in an unbiased manner, possible disease mechanisms for subsequent *in-vitro* and *ex-vivo* hypothesis driven experimentation.

Previous studies have suggested that low levels of vitamin D before or at onset of critical care increase risk of mortality (Braun et al., 2011; Braun et al., 2011) and 25(OH)D has a long half life (~2 months) (Mawer EB, 1971). This could imply that vitamin D levels have a causal relationship with disease susceptibility or with survival following intensive care. Consumption of vitamin D during acute states has been observed however (Bang et al., 2011; Reid et al., 2011). Replication in prospective cohorts and investigations in more homogenous groups of patients have not been carried out and are also necessary.

A possible explanation of the association between low levels of vitamin D and increased mortality is that prior low levels of vitamin D, acute consumption or both may increase susceptibility to infection, contribute to an altered immune and inflammatory response and increase disease severity and risk of death. Reverse causality, unaccounted confounders and other factors could equally explain these observations.

Further work is required in several areas. Investigations to ascertain the role of vitamin D in intensive care patients, particularly in those with severe sepsis and septic shock, as well as the value of vitamin D supplementation in these groups of patients are necessary. Studies to understand sepsis pathophysiology in better detail are also required to adequately classify patients and design effective interventions.

## Literature review conclusions

Studies have so far shown the relative importance of vitamin D in human health and disease. Molecules essential for its metabolism and function as well as the genomic binding sites of VDR are under positive selection. VDR and vitamin D enzymes are expressed in many organisms and in most human tissues. Both candidate approaches and genome-wide methods have recently shown its widespread and high-level functions in immune cells. Multiple observational studies show associations between vitamin D deficiency and autoimmune and inflammatory diseases.

Many limitations are present however and require careful consideration. Most experiments have been done in cancer cell lines or Epstein Barr Virus transformed B cells that are known to interact with vitamin D and VDR. Despite the technical challenges, more experiments using a variety of sources, sample types and stimulation conditions, in the case of gene expression profiling, are still required to fully understand the wealth of effects of vitamin D. The use of primary cells can bypass the limitations associated with cell immortalisation techniques and studies probing activated immune cells and the effect of vitamin D have not been performed extensively either.

The exact mechanisms as to how vitamin D exerts control and its molecular association to disease are not fully understood, in particular in relation to genetic variation and immune related disorders. Future studies could focus on important areas such as further regulatory functions of vitamin D and on its possible causal role and molecular mechanisms in disease. *In-vitro* and *ex-vivo* studies assessing DNA-protein

interactions and gene expression profiling can shed light on the functions and regulation of and by vitamin D.

In this thesis I explore several of these issues by using *in-vitro* and *in-vivo* techniques as well as investigating the importance and regulatory relationships of vitamin D in healthy individuals, in patients with MS and in patients with sepsis. The overarching thesis is that vitamin D is an important modulator of immunity and that its regulation, blood levels, actions and interaction with genetic variation are important in disease. I apply epidemiological and functional genomics approaches to show that vitamin D contributes importantly to immune and inflammatory responses and that vitamin D deficiency is an important risk factor in inflammatory conditions such as sepsis.

# CHAPTER 2: MATERIALS AND METHODS

## Systematic review of the literature

### SOURCES AND SEARCH LIMITS

Pubmed, Pubmed Central, the International Clinical Trials Registry Platform of the World Health Organisation, ClinicalTrials.gov and Google Scholar were used searching with no prior date and up until 28/September/2013; original investigations, review articles and trials, in humans or rodents, English language and with full text availability were included (except for key historical references).

### SEARCH STRATEGY

The criteria used to search consisted of the following: authors, institutions, and journals were not considered; searches were performed for terms appearing in the title, abstract or keywords using the terms below; articles were selected based on relevance by first searching in the title and abstract and then in the main content; reference lists of relevant articles were manually searched for additional citations not included above. A combination of the following search words and key terms were used for each topic:

- Vitamin D biology: vitamin D, vitamin D receptor, VDR, vitamin D biology, function, sources, metabolism, genetic determinants; bone, calcium homeostasis, immunity, immune system, VDR knockout, evolution; $1,25(OH)_2D_3$ –VDR/RXR complex, mechanism of action; genome-wide binding, expression, gene regulation, epigenetic regulation.

- Complex disease and vitamin D: complex disease, vitamin D and disease; rickets, type 1 diabetes, tuberculosis, hereditary vitamin D resistant rickets, multiple sclerosis, severe sepsis, septic shock; VDR polymorphisms; VDR mutations; genome-wide association studies; CYP24A1, CYP27B1, DHCR7, GC; susceptibility, prevention, treatment, supplementation; clinical trials, plasma levels, meta-analysis.

- Multiple sclerosis and vitamin D: multiple sclerosis, epidemiology, clinical presentation, diagnosis, susceptibility; clinical progression; multiple sclerosis aetiology, immunological basis, pathology; treatment, prognosis; epidemiology, vitamin D, multiple sclerosis, migration, month of birth, latitude; multiple sclerosis genetics, GWAS, vitamin D, VDR, CYP27B1, CYP24A1; vitamin D and EBV; vitamin D and interferon; vitamin D, multiple sclerosis, clinical trials, treatment, prevention.

- Sepsis and vitamin D: vitamin D, sepsis, severe sepsis, septic shock; definition, diagnosis, epidemiology; susceptibility, prognosis, genetics, clinical progression, clinical presentation; epidemiology, vitamin D, intensive care, sepsis, hospitalised patients, hospital; pathophysiology, sepsis aetiology, disease mechanisms; treatment, clinical trials; vitamin D, mechanisms, sepsis; vitamin D trials, sepsis, intensive care.

### ANALYSIS OF ENRICHMENT OF TRANSCRIPTION FACTOR BINDING IN DISEASE ASSOCIATED INTERVALS

For MS, GWAS regions were those as defined by (Sawcer et al., 2011) with a list provided by the NHGRI GWAS catalogue (http://www.genome.gov/gwastudies/; downloaded April 23, 2013). Disease variants were also as defined by NHGRI's GWAS catalogue for traits with 8 or more associated SNPs. A minimum number of SNPs was chosen arbitrarily to avoid false associations due to few observations. Disease intervals consisted of 150 kb bp added on each side of the reported GWAS variant. Distances for flanking regions were chosen to ensure that haplotypic blocks and distant regulatory elements such as enhancers were covered.

Analysis of overlap of genomic sets of intervals (disease intervals versus ChIP-seq binding lists to test for enrichment) was done using the Genome Association Tester as explained above (GAT, http://wwwfgu.anat.ox.ac.uk/~andreas/documentation/gat/contents.html) (Heger et al., 2013). GAT tests for non-random association between sets of genomic intervals by simulating within a genomic context and accounting for segment length distribution, chromosomal location and compositional biases while preserving local GC content. It computes the expected overlap and empirical p-value. Empirical false discovery rates (FDR) are computed by aggregating randomizations from multiple reference sets. GAT reports fold enrichment as the ratio of the observed over expected overlap. Settings used were nucleotide overlap, 10,000 permutations per test for P-value

calculation, ungapped contigs from the human genome hg19 track as background and significance considered at p-values <0.05 and q-values <0.05.

### ANALYSIS OF GENETIC VARIATION IN VDR BINDING SITES

*In silico* analysis was carried out using JASPAR CORE vertebrata database (http://jaspar.genereg.net/) with the RXRA::VDR (MA0074.1) Homo sapiens matrix model and a relative profile score threshold of 70%.

Vitamin D receptor tracks are those described in Ramagopalan et al. (2010). Overlapping intervals were tested using Galaxy tools (http://main.g2.bx.psu.edu/) based on SNPs and coordinates obtained from UCSC browser (http://genome.ucsc.edu/). VDR motifs were obtained using MEME-ChIP analysis (http://meme.sdsc.edu/meme4_6_1/intro.html).

Linkage disequilibrium SNPs were obtained from a proxy search in SNAP (http://www.broadinstitute.org/mpg/snap/ldsearch.php) based on 1000g pilot 1 with r^>0.8 for the CEU population and within 500,000 bp. Disease SNPs are those associated by GWAS and catalogued in NHGRI (http://www.genome.gov/gwastudies/ accessed 02/06/2011). SNPs and small indels in VDR binding intervals are based on SNP132. Copy number variants, insertions and deletions were based on the database of genomic variants and structural variation (DGV v10).

Computational analysis of predicted and experimental functional consequences of single nucleotide variants was done using Regulome (http://regulome.stanford.edu/index) (Boyle et al., 2012).

**ELECTROMOBILITY SHIFT ASSAYS TO ASSESS ALLELE SPECIFIC BINDING OF VDR AT DISEASE ASSOCIATED LOCI**

Electromobility shift assays (EMSA) were performed using standard protocols (Buratowski et al., 2001). Polyacrylamide gels at 5%, synthetic oligonucleotide probes and recombinant VDR-RXR-beta protein complexes were used. Annealed oligonucleotide probes were based on the sequences in Table 2 with the addition of 5' overhangs (agct).

**Table 2 Sequences for electromobility shift assays**

| |
|---|
| **>rs7859805 A [A/G]** |
| [agct]GTCAAGTCTGACTGGCCGTTGGGGGC[A/G]AAAGGTGTGACTGAAGC ACTGACAA |
| **>rs1154154 T [C/T]** |
| [agct]ATGGGGAAGTGGGTCCAATGGCCTCT[C/T]TGGCCTCTTAGTACAAA AAGTCTAT |

Sequences for EMSA probes with each allelic variant used to test VDR binding disruption.

Klenow enzyme (10 units, Fermentas EP0051) was used to radioactively label the resulting oligoduplexes with 0.74 Mbq of [$\propto^{32}$P] dCTP (Perkin Elmer, NEG513H250UC), 3mM of –dCTP and 75 $\mu$g of DNA per oligonucleotide sequence in a 20 $\mu$l mix per labelling reaction. Standard assays were carried out with 8 $\mu$g of radio-labelled oligoduplexes (0.0044 Mbq) in a 9 $\mu$l reaction containing 10 mM HEPES buffer (pH 7.8), 50 mM KCl, 1 mM EDTA, 1mM EGTA, 12.5% glycerol, 2 $\mu$g of each recombinant protein, 0.5 mg of poly dIdC (Sigma, P4929) and 0.5 $\mu$g of

bovine serum albumin and incubated at room temperature for 10 min. VDR full length recombinant protein (Invitrogen, P2190) and RXR-beta ligand binding domain recombinant protein (Invitrogen, PV4390) were used to bind the oligoduplex sequences. Protein-DNA complexes were pre-incubated at room temperature for 10 min before adding the radioactively labelled probe. Competition assays were carried out as standard with the addition of 10x and 100x of self, alternate allele or non-specific (a sequence not carrying the VDR DR3 motif) labelled and non-labelled oligoduplex probes accordingly (see main text for description of each lane for competition assays). For supershift analysis 1 $\mu$g per reaction of goat polyclonal IgG antibody against CTCF (Santa Cruz Biotechnology, sc-15914) and rabbit polyclonal IgG antibodies against NF$\kappa$B p65 (Santa Cruz Biotechnology, sc-109) and VDR (Santa Cruz Biotechnology, sc-1008) were used. The reaction was analysed by electrophoresis in a non-denaturing 5% polyacrylamide gel at 4°C in 1x TBE buffer. Gels were dried at 80°C in a vacuum and exposed over 12 hours on an autorad machine with high sensitivity photographic film (Carestream Kodak X-Omat LS film, Sigma, F1149).

## FAMILY BASED GENETIC ASSOCIATION TEST IN MULTIPLE SCLEROSIS

### *Ethical approval, recruitment, case ascertainment and data collection*

The Canadian Collaborative Project on the Genetic Susceptibility to MS (CCPGSMS) is a large family-based on-going cohort comprising MS patients and family members (Sadovnick et al., 1996). The CCPGSMS is a nation-wide Canadian collaboration of 15 MS clinics that uses structured questionnaires administered by trained interviewers to MS index cases, spouse controls and mothers of index cases and spouse controls

coupled to blood sampling for DNA based analysis. The CCPGSMS has been followed since 1972 starting in the London Multiple Sclerosis Clinic with updating of clinical definitions and an extended analysis of natural history since 1987 (Weinshenker et al., 1989; Weinshenker et al., 1987). The main objectives of the study are to carry out genetic epidemiological research and gather population-based data to understand the aetiology of MS (Sadovnick et al., 1996). The Universities of British Columbia and of Western Ontario approved the project with each clinic obtaining ethical approval from their institutional review board. All patients and family members provided informed consent.

***Sample processing, DNA extraction and storage***

Blood was collected from consenting patients and family members, stored at 4°C, -20°C or -80°C and DNA extracted using different methods that were standard at the time (Ebers et al., 1996). Duplicate samples are kept in London, Canada and in Oxford, UK for most patients and family members. A managed database using Microsoft Access is used to store clinical and demographic information.

***Sequenom iPlex genotyping and statistical analysis***

The study presented here included a total of 3,019 individuals, 1,858 female and 1,161 male, from 738 families with 1,360 individuals with definite MS and 1,659 of their unaffected first-degree relatives. Patients and family members were genotyped for 31 SNPs that were identified to be within vitamin D receptor binding sites and within 100 kb of MS susceptibility GWAS implicated genes (Table 3) (Ramagopalan et al., 2010; Sawcer et al., 2011). Total genomic DNA was extracted from whole blood as described above and genotyped by a core genomics service facility

(Wellcome Trust Centre for Human Genetics, Oxford; http://www.well.ox.ac.uk/ogc/home) using the Sequenom MassEXTEND protocol as per the manufacturer's instructions (http://www.sequenom.com; Sequenom, Inc.). Samples were discarded if low quality genotyping calls were made or consistent genotypes could not be obtained for a given individual as determined by the core facility. Pedigree structure and disease status were hidden for genotype generation. SNP alleles were tested for association with the transmission disequilibrium test (TDT) using PLINK (Purcell et al., 2007) with 10,000 permutations to correct for multiple testing. Chi-square tests were used for significance based on the number of times an allele is transmitted to affected offspring from heterozygous parents. To preserve linkage disequilibrium and linkage information between markers and siblings, permutations in PLINK constantly re-assign transmitted and un-transmitted status for all SNPs for a given family. Other members of the Ebers Lab had previously typed these samples for HLA-DRB1 alleles with either low- or high-resolution allele-specific PCR amplification (Lincoln et al., 2005). Significant genotypes were corrected for presence of HLA-DRB1*15.

**Table 3 Single nucleotide polymorphisms present within VDR binding sites used in a family based test of association to MS susceptibility.**

| SNP | Allele | Chromosome | Position | Poor genotyping rate |
|---|---|---|---|---|
| rs6695456 | C/T | 1 | 10270526 | X |
| rs56383778 | C/G | 1 | 117046977 | |
| rs59608162 | -/T | 1 | 200988378 | |
| rs12722635 | C/T | 10 | 6093140 | X |
| rs41295051 | A/G | 10 | 6111210 | |
| rs41295053 | A/G | 10 | 6111398 | |
| rs34080929 | -/G | 10 | 6114718 | |
| rs10905718 | A/G | 10 | 6114856 | |
| rs10747783 | C/T | 12 | 58176614 | |
| rs2071170 | C/G | 16 | 10970627 | |
| rs3087456 | A/G | 16 | 10970902 | |
| rs13335138 | A/G | 16 | 10985772 | X |
| rs7404786 | C/G | 16 | 11012550 | X |
| rs8057835 | C/T | 16 | 11012695 | X |
| rs8054198 | C/T | 16 | 11038360 | X |
| rs6498139 | A/G | 16 | 11047923 | |
| rs4781028 | C/G | 16 | 11059363 | |
| rs9302457 | A/G | 16 | 11059837 | X |
| rs35250210 | A/T | 16 | 11075543 | X |
| rs35494353 | -/T | 16 | 11146181 | |
| rs428530 | A/G | 16 | 85922859 | |
| rs305083 | A/G | 16 | 85935878 | |
| rs182511 | C/G | 16 | 85937656 | |
| rs17081891 | G/T | 18 | 67602921 | |
| rs17081900 | A/G | 18 | 67603002 | |
| rs6566448 | G/T | 18 | 67612527 | |
| rs17229961 | A/C | 18 | 67616613 | |
| rs280502 | G/T | 19 | 10491475 | |

| | | | | |
|---|---|---|---|---|
| rs6131014 | C/T | 20 | 44737383 | X |
| rs6074022 | C/T | 20 | 44740196 | X |
| rs2243115 | G/T | 3 | 159706280 | |

SNPs were tested using a family based association test, transmission disequilibrium, in the CCPGSMS. SNPs with poor genotyping rates were excluded from further analysis. Reference SNP alleles are reported as major/minor allele and -/ for indel based on dbSNP131.

# Integrative analysis of vitamin D in healthy individuals

### RECRUITMENT, SAMPLE COLLECTION AND CELL SEPARATION

Healthy volunteers were recruited and blood was collected and processed by other members of the J Knight group as described previously (Fairfax et al., 2012). Briefly, the Oxfordshire Research Ethics Committee (COREC reference 06/Q1605/55) granted ethical approval and informed consent was obtained from 288 healthy individuals. Whole blood was extracted using EDTA collection tubes (Vacutainer System, Becton Dickinson) and 50 ml of peripheral blood mononuclear cells (PBMCs) were purified using the Ficoll gradient method in order to remove polymorphonuclear cells. Cells were washed three times with Hanks-buffered saline solution (HBSS) without $Ca^{2+}$ or $Mg^{2+}$ (Invitrogen) to eliminate platelet contamination and cell counts were obtained using an automated cytometer (Countess). Purified cell populations were extracted by using magnetic beads (MACS method, Miltenyi) to positively separate $CD14^{+}$ and $CD19^{+}$ cells according to the manufacturer's instructions. Twenty million PBMCs were used as starting material for each cell type. Cells were then kept on ice and lysed immediately using RLT (Qiagen). Blood plasma was separated before Ficoll purification and stored immediately at -80°C.

### GENOMIC DNA AND RNA EXTRACTION

DNA and RNA extractions were carried out by the J Knight Lab (Fairfax et al., 2012) according to the manufacturer's instructions (Gentra Puregene Blood kit and RNeasy mini kit respectively, both Qiagen). For DNA, 1 to 2 ml of whole blood was used. Concentration measurements were done using the PicoGreen dsDNA assay

(Invitrogen) for DNA and Nanodrop and Bioanalyzer (Bioanalyzer RNA 6000 Nano, Agilent) for a subset of the total RNA extractions following manufacturers' instructions. cDNA conversions were carried using Invitrogen Superscript III System according to the manufacturer's instructions with oligo(dT) priming.

### GENE EXPRESSION MICROARRAY QUANTIFICATION AND ANALYSIS

Illumina HumanHT-12 v4 BeadChip gene expression array platforms were used for 288 samples as described by Fairfax et al. (2012). Briefly, after background subtraction (R packages lumi and limma) and removal of possible batch and array effects (principal component analysis), raw data were transformed and normalized (robust spline normalization method). A total of 29,022 probes for 286 B-cell and 287 monocyte samples remained after this. Array probes that had non-unique mapping, mapped to the sex chromosomes or mapped over SNPs (minor allele frequency >1%, 1000 Genomes Project, release 23/11/2010) were removed from further analysis. Dr. Julian Knight granted access to data.

### DNA GENOTYPING AND ANALYSIS

The study design focussed on candidate variants associated to vitamin D blood levels as reported by Wang et al. (2010) and Ahn et al. (2010). Data for these was extracted from the genome-wide genotyping performed by Fairfax et al. (2012). Dr. Julian Knight granted access to genotyping data. Briefly, DNA genotyping was carried out using the Illumina Human OmniExpress-12v1.0 BeadChips, NCBI36 Build, in order to map 733,202 genetic markers. After applying standard quality control filters 651,210 markers remained (SNP call rate > 96%; minor allele frequency (MAF) > 1%). Five individuals were removed after further quality control using multi-

dimensional scaling and identity-by-descent cluster analysis. The final analysis included 283 volunteers (161 female). The effect of population stratification on gene expression was assessed using the principal components of variance of genotyping. The number of probes associated to population stratification (n=65) did not differ to that expected by chance. PLINK (Purcell et al., 2007) was used for quality control and association analysis of the candidate vitamin D levels associated SNPs (see Table 1).

### ANALYSIS OF PLASMA VITAMIN D AND CORRELATION TO GENOTYPE AND TRANSCRIPTOME

Plasma metabolites were measured by a clinically and DEQAS (http://www.deqas.org/) approved laboratory using liquid chromatography mass spectrometry for albumin, C reactive protein (CRP) and 25-hydroxyvitamin $D_2$ and $D_3$ (25(OH)D) (Homerton Hospital Clinical Pathology Laboratory). Plasma 25(OH)$D_2$, 25(OH)$D_3$ and CRP were normalised to albumin and adjusted according to the dilution fraction as plasma samples were collected after mixing with HBSS. Whole blood is diluted in HBSS at a 2:1 ratio but extracted blood not always equals 50 ml as the process can present slight variation. Although individuals reported to be healthy, CRP was used to identify any with underlying inflammatory processes. All volunteers had values within the normal range according to the measuring laboratory's standard (<8). 25(OH)$D_2$ was used to determine if there were individuals taking ergocalciferol supplements but none were detected with values above 11 nmol/L and thus no additions were made to $D_3$ measurements. Quantitative trait and gene expression correlation analysis were carried out in collaboration with other members of the J Knight group (J. Radhakrishnan). Correlation of expressed genes and 25(OH)$D_3$

levels were done using Spearman's rank correlation in R. PLINK was used to perform the quantitative trait analysis between genotype and plasma 25(OH)D values (Purcell et al., 2007). Univariate and multivariate analysis was performed using SPSS (version 19). For pathway analysis Ingenuity Systems (IPA, build version 8.8) was used with the following settings: reference set whole genome, direct and indirect relationships, endogenous chemicals included and considering only molecules and relationships experimentally observed in humans.

## Clinical and molecular correlates of vitamin D and sepsis survival

### PATIENT RECRUITMENT AND SAMPLE COLLECTION

This study was carried out as part of the Genomic Advances in Sepsis project, an on-going collaboration of several UK intensive care units and research centres (GAinS, http://www.ukccg-gains.org/). Ethical approval for GAinS was obtained from the Central Office of Research and Ethics Committees (COREC) and the local ethics, research and development committees of participating centres.

Patients reported here were recruited from 2005 until 2011 from eight UK ICUs for the GAinS cohort. Adult patients who consented to participate in the study had been diagnosed with severe sepsis or septic shock as defined by Annane et al. (2005). Sepsis was due to community-acquired pneumonia (CAP) or faecal peritonitis (FP) and patients who were admitted to the ICU were enrolled. CAP was defined according to published diagnostic criteria (Angus et al., 2002) and FP was diagnosed at laparotomy and defined as inflammation of the serosal membrane that lines the abdominal cavity secondary to faecal contamination.

Consent was obtained from the patient or their legal representative and, where possible, was confirmed with surviving patients. Patients who did not consent, had pre-existing immune-suppression, were pregnant or under 18 years of age were excluded. Patients who were enrolled consented within 24 hours of admission to ICU and were followed up by telephone from hospital discharge until 6 months after discharge from ICU for survival data. One patient withdrew after surrogate consent and was therefore excluded from analysis. Patients were not enrolled consecutively;

reasons for not enrolling were not available at the time of analysis. All units subscribed to managing patients according to Surviving Sepsis Campaign Guidelines (Dellinger et al., 2013).

### OUTCOME DEFINITION FOR THE VITAMIN D IN GAINS STUDY

Primary outcomes were vitamin D deficiency prevalence on admission to ICU and association of vitamin D levels with death in-ICU, at 28 days, in-hospital and at 6 months post ICU admission. Twenty-six demographic and clinical variables including APACHE, SOFA and comorbidity data were collected for regression analysis. Secondary outcomes included molecular phenotyping using gene expression, correlation between gene expression and plasma molecules and network and pathway analysis.

### DATA SOURCES AND MEASUREMENT

Clinical and laboratory data were obtained from an online case report form completed and subjected to strict quality control by trained research nurses. Serial samples were obtained on days 1, 3 and 5 of ICU admission for DNA, RNA and plasma analysis. Peripheral blood leukocytes for RNA analysis were isolated at the bedside using the LeukoLock filter system (Ambion) within 5 minutes of collection (Idaghdour et al., 2008). Clinical data was collected from day 1 of admission to ICU. The hydroxylated form of plasma vitamin D (25(OH)D), plasma cathelicidin (LL-37) and leukocyte mRNA were extracted from the same samples.

RNA extractions were performed according to the manufacturer's instructions (Total RNA Isolation, Ambion). $25(OH)D_2$ and $25(OH)D_3$ were measured in undiluted

plasma samples using liquid chromatography tandem mass spectrometry by a DEQAS (http://www.deqas.org/) approved clinical laboratory (Homerton Hospital Clinical Pathology Laboratory). The lower detection limit is 10 nmol/L for both metabolites. Hybridisation of RNA was performed at the Sanger Institute using Illumina HT-12 v4 BeadChips and Illumina Human WG6v3 arrays.

An initial project by other members of the J Knight group (J Radhakrishnan) used a subset of the cohort presented here for gene expression analysis. A different platform was later chosen due to advances in technology and lowering costs. Microarray data in the work presented here was normalised across both platforms using the distance weighted discrimination (DWD) method and analysed using BeadStudio software (V3.1, Illumina) and Limma on R (v. 2.10.1, Bioconductor v. 2.5). The DWD method currently appears to be one of the most robust when normalising groups of different sizes and different platforms and additionally benefits from lower gene detection loss (Rudy et al., 2011). Pre-processing of array values was done using the lumi package on R (Du et al., 2008). Data was normalized after applying a variance stabilizing transformation. Outliers and probes with low detection confidence were removed before further analysis. Normalisation and gene expression analysis were carried out in collaboration with other members of the J Knight group (P. Humburg).

Plasma cathelicidin measurements were carried out using a commercial enzyme-linked immunosorbent assay (Hycult Biotech, HK321 human LL-37 kit).

**UNIVARIATE AND BINARY LOGISTIC REGRESSION ANALYSES**

The distribution of plasma 25(OH)D and cathelicidin was analysed and pair-wise comparisons were performed using non-parametric tests in SPSS (v 20). Binary logistic regression and Cox regression analysis were done to predict all-cause mortality following an episode of severe sepsis or septic shock (SPSS, v 20).

For the vitamin $D_2$ form ($25(OH)D_2$) measured by tandem mass spectrometry, >95% of patients had plasma levels below the detection limit and were thus not taken into account. The sum of $25(OH)D_2$ and $25(OH)D_3$ was used throughout analyses although pair-wise mortality tests were not materially modified when only using $25(OH)D_3$. The term 25(OH)D is used to indicate the sum of both vitamin D forms in this study.

Exact two tailed p-values were obtained with significance considered at <0.05 using Mann Whitney U tests. Pair-wise analysis was carried out for vitamin D levels according to mortality (in-ICU, at 28 days, in-hospital and at 6 months), season (winter plus spring vs. summer plus autumn), diagnosis (CAP vs. FP), gender, and age (divided by the median: ≤64 years vs. >64 years). Binary logistic regression was carried out with predictors considered significant if present at both forward and backward stepwise conditional models: probability for entry <0.05 and for removal >0.1 and Hosmer and Lemeshow goodness of fit tests >0.05.

The Cox regression model was run in the same way as binary logistic regression on the unadjusted dataset. Season of ICU admission was excluded from the main model as it was considered to interact with vitamin D, decreased the goodness of fit test (to <0.05) and was not significant when included.

Potential confounding variables for vitamin D measurements that were accounted for included age, gender, ICU length of stay, diagnosis, acute shifts of 25(OHD)D plasma levels, day of sampling and season. A potential confounder for measurements of plasma 25(OH)D levels in the critically ill is haemodilution. Data for this is equivocal (Cecchi et al., 2011; Krishnan et al., 2010) and although whole body half-life of vitamin D is long (~2 months) (Jones, 2008) there are indications that vitamin D could be consumed during the acute phase response (Bang et al., 2011; Bertoldo et al., 2010; Reid et al., 2011). This was addressed through serial blood measurements for days 1, 3 and 5 of ICU admission and found that measurements from samples at all time points showed similar significant differences in survival at 6 months.

Vitamin D measurements were performed in a clinically and DEQAS accredited laboratory in a single batch using the gold standard method. Variables with missing data for more than 10% of patients were excluded from the analysis. Two patients were lost to follow-up for mortality at 6 months and were treated as missing data.

### MICROARRAY GENE EXPRESSION MEASUREMENTS, DIFFERENTIAL GENE EXPRESSION ANALYSIS AND PLASMA-TRANSCRIPTOME CORRELATIONS

The leukocyte whole genome transcriptome was analysed considering genes differentially expressed by comparing survivors and non-survivors at a minimal log fold change of 0.5 and a false discovery rate of <0.05 using R packages on Bioconductor (Limma v. 2.10.1, Bioconductor v. 2.5). T tests were used to identify differentially expressed genes with significance <0.05. Differentially expressed genes were correlated to plasma levels of several molecules as monotonic functions using

Spearman's rank correlation coefficient in R. Significance was considered with a T-test p-value <0.05 at various false discovery rates as presented in the main text.

Univariate, multivariate and regression models were performed using SPSS (v 20) as described above. Heatmaps were plotted based on unsupervised hierarchical clustering using the Euclidean distance and sorted according to correlation with polynucleocyte count for expressed genes that were correlated to both plasma 25(OH)D and polynucleocyte count on the vertical axis and patient survival on the horizontal axis. Z-scores were used to represent gene expression values standardised across samples and relative to the mean.

#### ENRICHMENT TESTS, CLUSTER AND PATHWAY ANALYSES

Enrichment tests were carried out using the genomic association tester (GAT) as described above and by Heger et al. (2013). Segment overlap was tested using all expressed genes in our cohort as background with 100,000 permutations per test. Transcriptional start and end sites were extended by 50 kb for expressed genes and genome-wide binding peaks were used as reported by the original studies. Significance was considered at p-values <0.05 and q-values <0.05.

For pathway analysis Ingenuity Systems (IPA, build version 192063 and content version 14400082) was used with the following settings: reference set HumanWG-6 v3.0, direct and indirect relationships, endogenous chemicals included and considering only molecules and relationships experimentally observed in humans.

# CHAPTER 3: MOLECULAR CHARACTERISATION OF GENETIC REGULATORY VARIANTS AND VITAMIN D RECEPTOR BINDING IN AUTOIMMUNE DISEASE

## Introduction

The importance of vitamin D as a general immune modulator has emerged in the last several years (Holick, 2007). Evidence of vitamin D being causal in autoimmune conditions is growing and its value as a therapeutic agent is currently being tested. MS is an autoimmune disease where several observations point to vitamin D as a causal agent (Ascherio et al., 2012). Other diseases have weaker evidence however.

The presence of genetic variation within transcription factor binding sites may alter binding by disrupting the DNA sequence necessary for recognition of the DNA binding domain on transcription factor proteins (Reddy et al., 2012). Disrupted binding may have downstream effects in gene expression and contribute to disease (Knight et al., 2004).

The first genome-wide map of VDR binding was recently described by our group using chromatin immunoprecipitation (ChIP) followed by massively parallel DNA sequencing (ChIP-seq) (Ramagopalan et al., 2010). This map allowed a detailed and unbiased genome-wide view of VDR binding in B cells and can serve as a basis to evaluate binding and genetic variation. Polymorphisms that disrupt transcription factor binding have previously been shown to be important in disease but only one has been described for VDR (Knight et al., 2004; Knight et al., 1999; Ramagopalan et al., 2009).

In this chapter I investigate the impact of disease-associated genetic variation on VDR binding. I focus on candidate methods using molecular genetics techniques and complement this with computational genome-wide analyses. These support the findings at specific loci and extend the notion that vitamin D has broad effects at disease relevant regions genome-wide. Each method has important limitations but taken together provide a framework that show examples of the impact of genetic variation at disease associated loci. Although I provide results that implicate vitamin D in other autoimmune conditions, I use multiple sclerosis as an example and motivation for the experiments performed in order to understand vitamin D receptor binding and disease associations.

I first hypothesised that VDR binding would occur near genetic disease associated signals more frequently than expected by chance and that disease associated genetic variation found within VDR motifs would disrupt binding. Using the VDR ChIP-seq map I first tested for binding enrichment near genome-wide association signals. I then searched for genetic variation associated to disease present in VDR binding intervals and motifs across the genome and finally carried out in vitro experiments to validate bioinformatics predictions.

The current resolution of ChIP-seq depends on a DNA shearing step and is usually greater than 300-400 bp (Park, 2009). Transcription factor motifs typically span less than 20 bp and, in the case of VDR, the one most commonly found is 15 bp long (DR3; see Figure 14 in results section below). Variation both within and outside of motifs is thought to be important in transcription factor function (Reddy et al., 2012).

Recent studies have suggested that most allelic occupancy results from variation outside of the DNA binding motif and that transcription factor occupancy up to 100 kb from the transcriptional start site may contribute to gene expression (Reddy et al., 2012).

Given the previous observations I hypothesized that SNPs present within binding sites, although not necessarily within identified canonical motifs, may still be directly relevant to disease susceptibility. Evidence associating vitamin D, genetic variation and susceptibility to MS is increasing (Ramagopalan et al., 2011; Sawcer et al., 2011) and suggest that disruption of vitamin D receptor binding to its genome targets may be important (Ramagopalan et al., 2009). Considering the above I performed a focused search for variants near MS associated genes using the ChIP-seq map described previously. I then tested the association to susceptibility of SNPs found within VDR binding sites near MS GWAS implicated genes in a Canadian cohort of MS families.

In summary, in this chapter I address the relationship between genetic variation, VDR binding and disease susceptibility. I find that VDR binds near disease associated genetic variation more than expected by chance, that autoimmune disease associated variation disrupts VDR binding and that a variant within a VDR binding site is associated to MS susceptibility risk.

***Molecular characterisation of genetic regulatory variants and vitamin D receptor binding in autoimmune disease:***

- Is VDR binding enriched near disease-associated genomic intervals?
- Are disease associated single nucleotide polymorphisms present within VDR motifs? Do these modulate VDR binding?
- Are genetic variants found within VDR binding intervals near MS-associated genes? Are these variants associated to multiple sclerosis susceptibility?

**Figure 11 Questions addressed in chapter three.**

# Results

## I. GENETIC VARIATION WITHIN VITAMIN D RECEPTOR BINDING MOTIFS

### *VDR binding is enriched near disease associated variants*

Using a recent update of all GWAS reported to date I tested enriched VDR binding near trait associated genetic variants (http://www.genome.gov/gwastudies/; accessed April 23, 2013)(Hindorff et al., 2009). Non-random overlap of VDR binding and disease intervals (150 kb on each side of the main reported associated SNPs) using ChIP-seq regions was investigated using the Genomic Association Tester (GAT) tool (Heger et al., 2013) as described in chapter two. I tested 332 traits from GWAS that had 8 or more independently associated SNPs. Results shown in Figure 12 are traits associated to autoimmunity. In total, 71 traits had significant values (p-value <0.05 and q-value <0.05) (for full list see appendix Table 22).

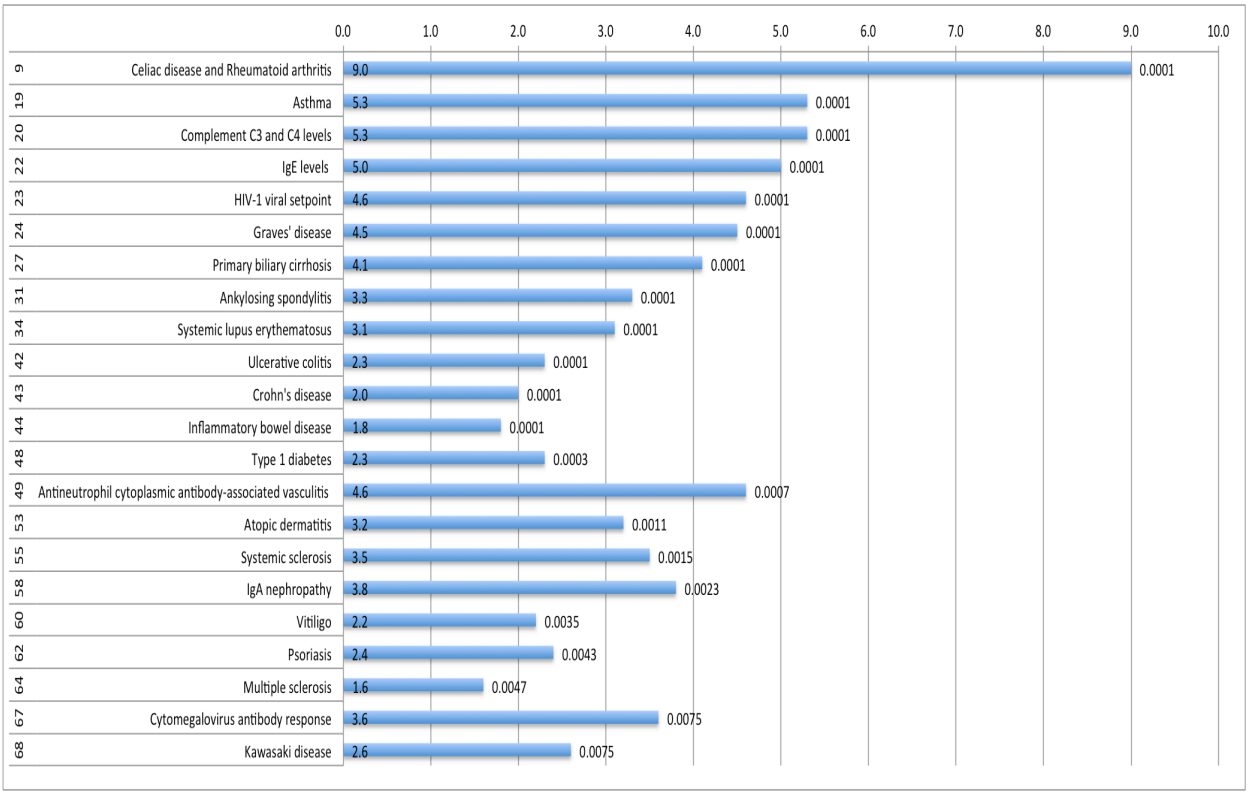



**Figure 12 Immune traits that are significantly enriched for VDR binding.**

Traits shown are significantly enriched for vitamin D receptor binding with p- and q-values below 0.05. Ranks (left), fold-change (middle) and q-values (right) are shown. Results are sorted and ranked according to q-value followed by fold change. Non-autoimmune traits that were significant are shown in Table 22.

As a comparison I tested androgen receptor (AR) binding in the same way for the 332 traits and found far fewer significant associations (Figure 13, all significant and marginal associations shown). The most significant associations were with renal function related traits. Gender differences and a role for sexual hormones in chronic kidney disease incidence and outcome have been well documented although

underlying mechanisms are not understood (Kummer et al., 2012). Prostate cancer was marginally non-significantly associated (corrected p-value=0.06).

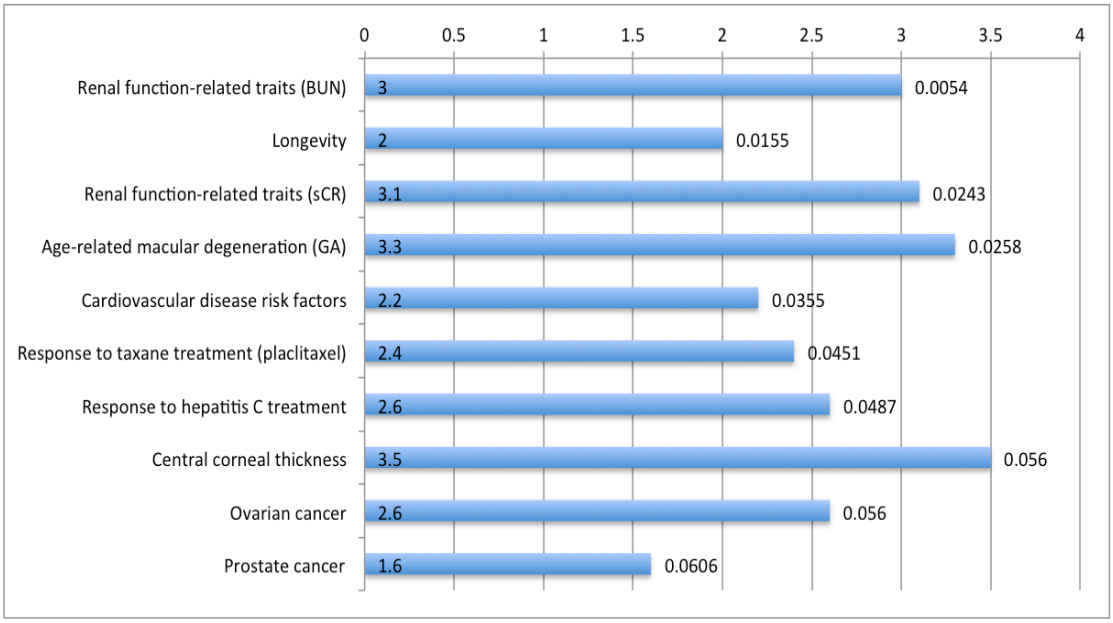



**Figure 13 Analysis of overlap between genome-wide AR binding and GWAS disease intervals.**

Traits significantly enriched for AR binding with p- and q-values below 0.05 plus marginal associations (q-value 0.06) are shown. Fold-change (left) and q-values (right) are shown. Results are sorted and ranked according to q-value followed by fold change. All significant results are shown. Data for AR binding was obtained from Yu et al. (2010).

***Disease associated polymorphisms disrupt VDR motifs***

I next analysed whether SNPs associated with disease by GWAS could be found within VDR ChIP-seq intervals and predicted motifs. I performed this for 523 traits and found that there are low levels of genomic variation within binding intervals (as

defined by ChIP-seq results and mapping tools available at the time of analysis for ChIP-seq) and within VDR binding motifs (as defined by *in-silico* analysis of ChIP-seq intervals based on DR3 and DR4 type motifs). I analysed common SNPs and disease associated SNPs as well as structural variants (Table 4). I found that relatively few SNPs are present within intervals but identified two examples of SNPs within VDR binding motifs in strong linkage disequilibrium with disease-associated polymorphisms.

**Table 4 Genetic variation found within vitamin D receptor motifs and binding sites.**

| | **Within motifs** | **Within binding intervals** |
|---|---|---|
| Number of sites | 1311 | 2776 |
| Number of base pairs | 19665 | 2124149 |
| Total number of SNPs | 153 | 13166 |
| Disease associated SNPs | 0 | 21 |
| SNPs in LD with disease associated variants | 2 | 130 |
| Number of indels | 2 | 42 |
| Number of inversions | 12 | 52 |
| Number of CNVs | 421 | 1456 |
| Number of inversion break-points | 0 | 1 |

SNPs and coordinates were obtained from the UCSC browser (http://genome.ucsc.edu/) for ChIP-seq intervals as defined by Ramagopalan et al. (2010). VDR motifs were obtained using MEME-ChIP analysis (http://meme.sdsc.edu/meme4_6_1/intro.html). Linkage disequilibrium SNPs were obtained from a proxy search in SNAP

(http://www.broadinstitute.org/mpg/snap/ldsearch.php) based on 1000g pilot 1 with r^>0.8 for the CEU population and within 500,000 bp. Disease SNPs are those associated by GWAS and catalogued in NHGRI (http://www.genome.gov/gwastudies/ accessed 02/06/2011). SNPs and small indels in VDR binding intervals are based on SNP132. Copy number variants, insertions and deletions were based on the database of genomic variants: structural variation (DGV v10).

***In silico analysis of disease associated polymorphisms within VDR motifs***

Vitamin D response elements commonly possess either a direct repeat of two hexanucleotide half-elements with a spacer of three nucleotides (DR3) or an everted repeat with six nucleotides (ER6) (Haussler et al., 2011). DR3 is the most common and a DR4 type (spacer with 4 nucleotide elements) as well as several others have been described (Pike et al., 2010). The logo sequence below, derived from the ChIP-seq genome-wide map of VDR binding in B cells, represents the most common motif found, a combination of DR3 and DR4 types (Figure 14). The two SNPs in high linkage disequilibrium with variants associated with autoimmune diseases were found at conserved sites within VDR motifs (Table 5).

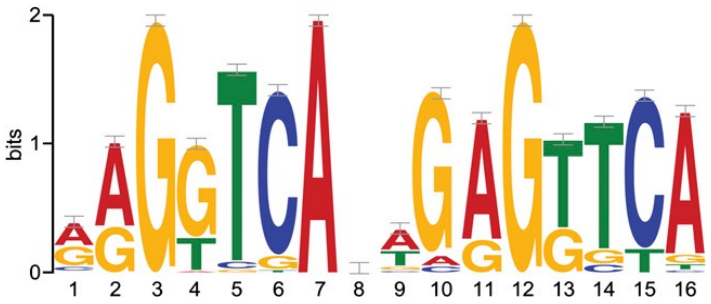



**Figure 14 Vitamin D receptor consensus motif.**

Canonical VDR binding site as defined in (Ramagopalan et al., 2010). The identified disease associated variants found within VDR motifs occupy the 6th and 9th nucleotide positions for the DR3 type motif and the 7th and 10th positions for the DR4 type motifs (rs7859805 and rs1154154 respectively).

These SNPs are in high linkage disequilibrium (r^2>0.8) with GWAS variants of autoimmune diseases, coeliac disease (CD), rheumatoid arthritis (RA) and narcolepsy. In the previous analysis of overlap of disease genomic intervals and VDR binding sites I showed that CD and RA are significantly enriched with a fold change of 9 (observed over expected ratio, p-value=0.0001, q-value=0.0009). Narcolepsy was not tested, as fewer than 8 polymorphisms have been associated to it.

**Table 5 Single nucleotide polymorphisms associated to disease found within VDR binding motifs.**

| **GWAS SNP (-risk allele)** | rs1953126-T | rs1154155-G |
|---|---|---|
| **Proxy SNP - minor allele (r^2, D')** | rs7859805-A (0.933, 0.966) | rs1154154-C (0.932, 1) |
| **Chromosome** | 9 | 14 |
| **Chromosome start** | 122703944 | 22072112 |
| **Chromosome end** | 122703945 | 22072113 |
| **Distance, bp** | 23623 | 412 |
| **Trait** | CD and RA | Narcolepsy |
| **Nearest gene** | TRAF1 | TCRA |
| **Number of motifs (strand)** | 3 (+) | 5 (-), 1 (+) |

*TRAF1* = tumour necrosis factor receptor associated factor 1; *TCRA* = T cell receptor alpha chain. Risk allele is also the minor allele. CD=coeliac disease; RA=rheumatoid arthritis. The predicted motif with disease associated SNPs near *TRAF1* is on the positive strand while on the negative for *TCRA*.

The SNP rs7859805, present within a vitamin D receptor motif in the promoter region of *TRAF1* is in near perfect LD (r^2 =0.933) with rs1953126, associated to coeliac disease and rheumatoid arthritis. The SNP rs1154154, also found within a vitamin D receptor motif in a gene dense region, is in near perfect LD (r^2 =0.932) with rs1154155, associated to narcolepsy. I also observed multiple VDR motifs near these disease associated SNPs (Table 5).

Figure 15 shows the genomic location and context of the SNPs rs7859805 and rs1154154. The two SNPs are located in potential regulatory elements as defined by DNase hypersensitivity mapping using data from the ENCODE Project suggesting a

putative functional role for this allelic variation (Dunham et al., 2012). The disease associated SNPs found within VDR motifs occupy the 6th and 9th positions for the DR3 type motif and the 7th and 10th positions for the DR4 type motifs (rs7859805, TRAF1, and rs1154154, TCRA, respectively) (Figure 14).

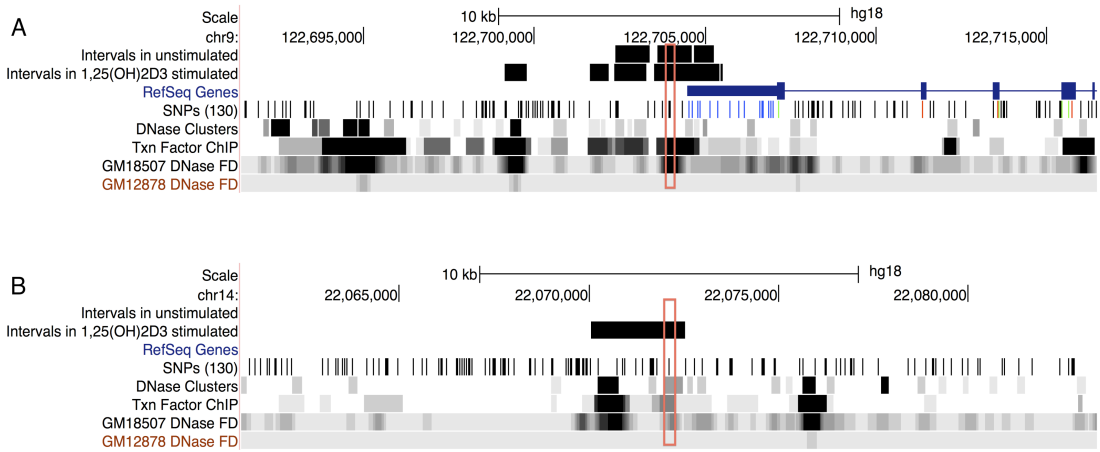



**Figure 15 Genomic locations of (A) rs7859805, associated to coeliac disease and rheumatoid arthritis, and (B) rs1154154, associated to narcolepsy.**

The first two tracks show genome-wide binding maps of VDR in two lymphoblastoid Hapmap cell lines (GM10855 and GM10861, statistically significant peaks called as intervals in unstimulated and 1,25$(OH)_2D_3$ stimulated cells) (Ramagopalan et al., 2010). Genes (Refseq; *TRAF1* in A), common SNPs (dbSNP 130), DNase I hypersensitivity clusters from diverse cell lines (ENCODE), transcription factor binding regions (Txn Factor ChIP, ENCODE) and DNase I hypersensitivity density tracks for two lymphoblastoid cell lines (GM18507 and GM12878 DNase FD ENCODE) are shown.

To better understand how genetic variation at these sites may cause disruption I analysed *in-silico* predictions of VDR binding with JASPAR (Sandelin et al., 2004) based on the DR3 type. Known transcription factor binding sites can be described quantitatively by position weight matrix (PWM) (Sandelin et al., 2004). PWM profiles are generated from experimental data by calculating the probability of observing a certain nucleotide sequence. A putative binding site can be compared to this PWM and a predictive score is derived. For rs7859805 (*TRAF1*), with a minimum threshold of 70%, I found a VDR binding site predicted for the minor (A) (relative score 0.72) but not for the major allele (G). For rs1154154 (*TCRA*), the motif is found on the negative strand and at a 70% threshold a site is only predicted for the major allele (T) (relative score 0.74) (appendix Table 23).

***In vitro evidence of allele specific binding of VDR rs7859805 and rs1154154***

Bioinformatic analysis predicted allele specific binding (ASB) of VDR at rs7859805 and rs1154154. Both SNPs lie within putative regulatory regions and are of potential disease relevance based on high linkage disequilibrium with GWAS SNPs for autoimmune conditions. In order to experimentally assess allele specific binding at these variants I used electromobility shift assays (EMSA). EMSAs remain the current standard to determine the effect of genetic variation on DNA-protein interactions despite their important limitations.

EMSAs consist of polyacrylamide gels where synthetic nucleotide probes interact with specific transcription factors of interest. Radioactive cytosine is incorporated into the DNA sequence to allow visualisation of DNA-protein interactions. A single

nucleotide change using two sequences tested independently allows qualitative assessment of the specificity of protein binding to DNA. Figure 16 and Figure 17 show the results of EMSA experiments interrogating how binding may be modulated by rs7859805 and rs1154154 using VDR-RXR-beta recombinant proteins.

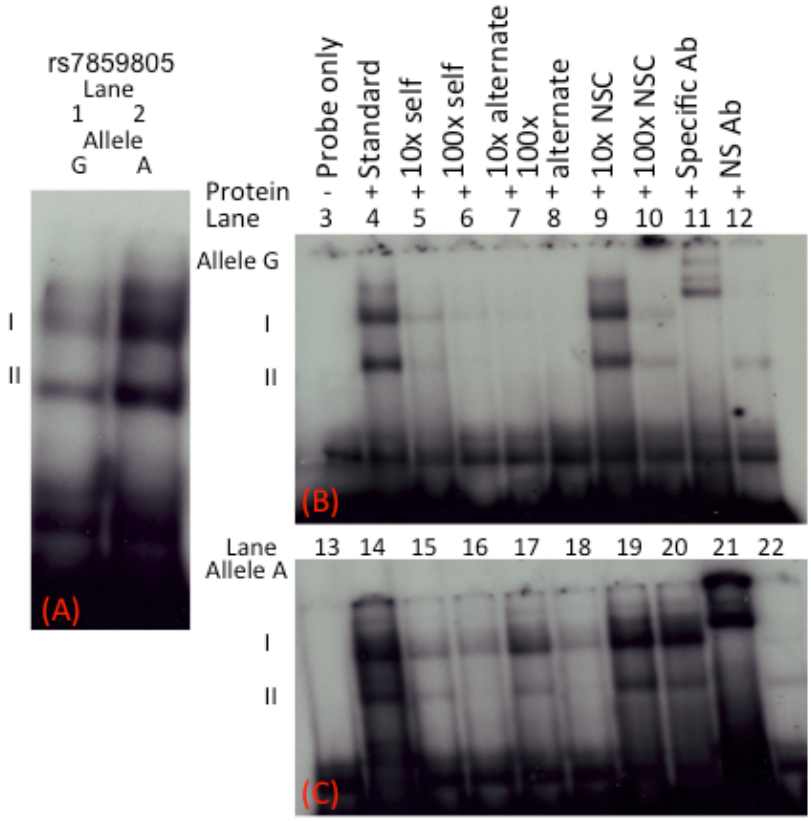



**Figure 16 Electromobility shift assay investigating VDR binding at rs7859805, near *TRAF1* and implicated in coeliac disease and rheumatoid arthritis susceptibility.**

(A) Comparison of G and A alleles (lanes 1 and 2) incubated with recombinant VDR-RXR demonstrates a number of retarded complexes. Analysis of binding using radiolabelled allele G (panel B) or allele A (panel C). Probe alone (lane 3, 13) or incubated with recombinant VDR-RXR (lanes 4-12, 14-22) shown. Complexes I and II are shown to be competed using 10x or 100x molar excess of unlabelled self oligoduplex (lanes 5 and 6, 15 and 16) or the unlabelled alternate allele (lanes 7 and 8, 17 and 18). Competition using an unlabelled oligoduplex corresponding to an unrelated DNA

sequence shown (lanes 9 and 10, 19 and 20). Antibody studies using anti-VDR (lanes 11, 21) or unrelated antibody (lanes 12, 22; anti-CTCF). Gel representative of four independent experiments. NSC, non-specific.

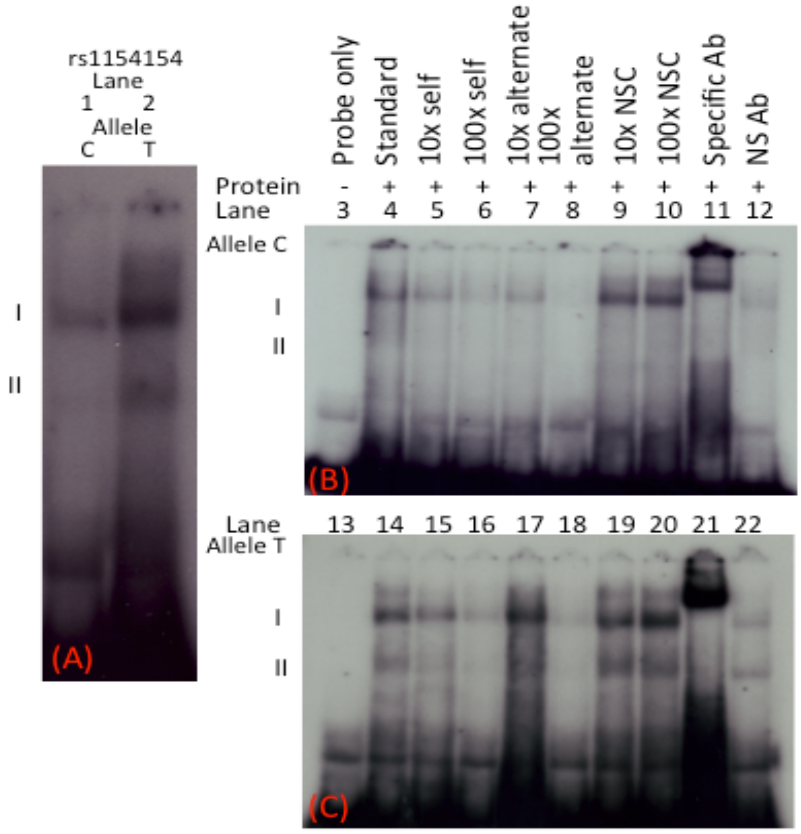



**Figure 17 Electromobility shift assay of VDR binding at rs1154154, near *TCRA* and implicated in narcolepsy susceptibility.**

(A) Comparison of C and T alleles (lanes 1 and 2) incubated with recombinant VDR-RXR demonstrates a number of retarded complexes. Analysis was carried out using radiolabelled allele C (panel B) or allele T (panel C). Probe alone, recombinant VDR-RXR incubation, specific and non-specific allelic competition and antibody analysis was carried out as for

Figure 16 with corresponding lane numbers. Two complexes (I and II) are identified. Unrelated antibody was anti- NFκB p65. Gel representative of four independent experiments. NSC, non-specific.

For rs7859805, there was evidence of two specific retarded complexes showing higher binding affinity for the A allele (Figure 16). Addition of 10x and 100x molar excess of unlabelled oligoduplexes corresponding to the same or alternate allele show weakening or disappearance of the DNA-protein band due to competition for the protein between labelled (radioactive) and unlabelled sequences. The results are consistent with the A allele outcompeting the G allele for rs7859805. Non-specific competition with an unlabelled oligoduplex sequence not thought to bind VDR does not alter the strength of the bands expect at high concentrations with allele G for rs7859805. The addition of anti VDR antibody causes a super-shift consistent with VDR binding in these complexes. A control antibody to NFkB p65, in lane 10 does not cause shifted bands but does interfere with binding resulting in loss of the complex.

Allele-specific binding was also investigated for rs1154154. Here there was evidence of one specific retarded complex showing higher binding affinity for the T allele (Figure 17). Competition EMSA demonstrates the T allele outcompeting the C allele. Non-specific competition with an unlabelled oligoduplex sequence not thought to bind VDR does not alter the strength of the bands. The addition of anti VDR antibody causes a super-shift consistent with VDR binding in this complex and as previously, a

control antibody to NFkB p65 does not cause shifted bands but does interfere with binding resulting in loss of the complex.

***Bioinformatics analysis of the functional impact of rs7859805 and rs1154154***

In order to understand known and predicted functional impact of rs7859805 and rs1154154 I first looked at ENCODE and eQTL data. As shown in Figure 15 both SNPs lie within regulatory regions (DNase tracks) and where transcription factors other than VDR also bind (Txn Factor ChIP track), indicating likely variant functionality. A recently published tool, RegulomeDB, combines known experimental data (ENCODE) and predicted functional consequences based on alteration of transcription factor binding to provide a score of functional likelihood (Boyle et al., 2012). It does not include vitamin D receptor data however.

High scores and many sources of experimental evidence of functional importance resulted for both variants using this (see appendix Table 24). For rs7859805, it determined that it would likely affect transcription factor binding and is known to act as an eQTL for PHD finger protein 19 (*PHF19*) in lymphoblastoid cells (Table 25; score 1b where the highest level of evidence is 1a, see Table 24). Additionally, NFKB1 binds over rs7859805 and a motif of pleomorphic adenoma gene 1 (PLAG1) is predicted to be disrupted by this polymorphism. Open chromatin and several histone modifications that identify regulatory regions are found covering rs7859805 in lymphoblastoid cells (Table 25).

In the case of rs1154154 the score for functional association was 3a, indicating less available experimental evidence (Table 26). Experimental data shows that rs1154154

lies within open chromatin and regulatory regions marked by histone modifications in lymphoblastoid cells. SPI1, BCLAF1 and POLR2A bind over it and predicted motifs for SRF and PLAG1 are disrupted. Further experimentation is required to assess whether rs1154154 has additional impact such as acting as an eQTL.

## II. GENETIC VARIATION ASSOCIATED TO MULTIPLE SCLEROSIS PRESENT WITHIN VITAMIN D RECEPTOR BINDING SITES

### *Identification of genetic variation within VDR binding sites near MS-associated genes*

In this experiment I tested the association of genetic variants lying within VDR binding sites and MS susceptibility risk by performing a family based association test (Spielman et al., 1996). I hypothesised that SNPs located within VDR binding sites would have distorted transmission in MS affected families. I first screened for variation using results from the ChIP-seq map of VDR binding in B cells as presented in the previous section. This analysis identified common variation and disease associated SNPs within binding sites although none previously linked to MS. I present the results of bioinformatics analyses of these variants above (Table 4). These variants were not found within canonical VDR motifs although previous studies indicate that variants near motifs remain functionally relevant (Reddy et al., 2012). In order to establish the relevance to disease of common variants within VDR binding sites I tested for the association of these SNPs with MS susceptibility in a Canadian cohort of affected families.

### *The Canadian MS cohort*

The Canadian Collaborative Project on the Genetic Susceptibility to MS (CCPGSMS) is a large family-based cohort comprising MS patients and family members (Sadovnick et al., 1996). It is a nation-wide collaboration that studies the epidemiology and genetics of MS aetiology and has been ongoing since 1972 with updating of clinical definitions and high disease ascertainment (Scalfari et al., 2010). Ethical approval and informed consent were provided for this study. The cohort

comprised a total of 3019 individuals, 1858 female and 1161 male, from 738 families with 1360 individuals with definite MS and 1659 of their unaffected first-degree relatives. Sex ratios in this cohort were 1.6 in all members (1867:1164, female to male); 2.4 in affected individuals (986:374) and 1.1 in unaffected individuals (872:787), keeping with the general trends observed in other cohorts (Kingwell et al., 2013).

***A SNP in a VDR interval is associated to MS susceptibility***

The cohort was genotyped for 31 SNPs found within VDR binding intervals near MS GWAS associated SNPs (see Table 3 in chapter two). Variants tested in the previous section, rs7859805 and rs1154154, are common SNPs that have not been previously associated with MS susceptibility or are near MS implicated genes and were thus not tested here. After quality filters, 21 SNPs remained for transmission disequilibrium testing (TDT) (Table 6) (see Chapter 2: Materials and Methods). The TDT uses chi-squared significance assessment based on the number of times that an allele is transmitted to affected offspring by parents that are heterozygous. PLINK was used for statistical analysis (Purcell et al., 2007) with 10,000 permutations to correct for multiple testing. After multiple testing correction, a variant 6877 bases upstream of the transcriptional start site of *IL2RA*, rs41295051, remained significant (transmitted 24, un-transmitted 51, OR 0.47, corrected p-value 0.022) (Table 6).

**Table 6 Family based test of genetic association of SNPs found within VDR binding sites and within or near MS associated genes.**

**Transmission disequilibrium test**

| CHR | SNP | BP | Nearest gene | A1 | A2 | T | U | OR | CHISQ | P | P corrected |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 | rs56383778 | 117046977 | CD58 | C | G | 2 | 3 | 0.667 | 0.200 | 0.655 | 1.000 |
| 1 | rs59608162 | 200988378 | KIF21B | - | - | 5 | 6 | 0.833 | 0.091 | 0.763 | 1.000 |
| 3 | rs2243115 | 159706280 | IL12A | G | T | 112 | 123 | 0.911 | 0.515 | 0.473 | 1.000 |
| 10 | rs41295051 | 6111210 | IL2RA | A | G | 24 | 51 | 0.471 | 9.720 | 0.002 | 0.022 |
| 10 | rs41295053 | 6111398 | IL2RA | G | A | 0 | 2 | 0.000 | 2.000 | 0.157 | 0.937 |
| 10 | rs34080929 | 6114718 | IL2RA | - | - | 0 | 0 | NA | NA | NA | NA |
| 10 | rs10905718 | 6114856 | IL2RA | G | A | 292 | 277 | 1.054 | 0.395 | 0.530 | 1.000 |
| 12 | rs10747783 | 58176614 | TSFM | - | C | 0 | 0 | NA | NA | NA | NA |
| 16 | rs2071170 | 10970627 | CIITA | C | G | 18 | 23 | 0.783 | 0.610 | 0.435 | 1.000 |
| 16 | rs3087456 | 10970902 | CIITA | G | A | 240 | 223 | 1.076 | 0.624 | 0.430 | 0.999 |
| 16 | rs6498139 | 11047923 | CLEC16A | A | G | 98 | 99 | 0.990 | 0.005 | 0.943 | 1.000 |
| 16 | rs4781028 | 11059363 | CLEC16A | - | C | 0 | 0 | NA | NA | NA | NA |
| 16 | rs35494353 | 11146181 | CLEC16A | - | - | 0 | 0 | NA | NA | NA | NA |
| 16 | rs428530 | 85922859 | IRF8 | G | A | 36 | 49 | 0.735 | 1.988 | 0.159 | 0.938 |
| 16 | rs305083 | 85935878 | IRF8 | G | A | 201 | 214 | 0.939 | 0.407 | 0.523 | 1.000 |
| 16 | rs182511 | 85937656 | IRF8 | G | C | 30 | 36 | 0.833 | 0.546 | 0.460 | 1.000 |
| 18 | rs17081891 | 67602921 | CD226 | - | T | 0 | 0 | NA | NA | NA | NA |
| 18 | rs17081900 | 67603002 | CD226 | - | A | 0 | 0 | NA | NA | NA | NA |
| 18 | rs6566448 | 67612527 | CD226 | - | T | 0 | 0 | NA | NA | NA | NA |
| 18 | rs17229961 | 67616613 | CD226 | A | C | 40 | 39 | 1.026 | 0.013 | 0.910 | 1.000 |
| 19 | rs280502 | 10491475 | TYK2 | T | G | 174 | 172 | 1.012 | 0.012 | 0.914 | 1.000 |

Family based test of genetic association of SNPs found within VDR binding sites and within or near MS associated genes. Corrected p-values are after multiple testing with 10,000 permutations. CHR = chromosome number, SNP = SNP identifier, A1 = minor allele code, A2 = major allele code, T = transmitted minor allele count, U = un-transmitted allele count, OR = TDT odds ratio, CHISQ = TDT chi-square statistic, P = TDT asymptotic p-value.

*HLA-DRB1*1501* carries the highest risk for MS susceptibility with reported ORs of 3.1 for heterozygous individuals (Sawcer et al., 2011). Patients and family members studied here had been previously genotyped at the MHC II locus with allele-specific PCR amplification (Lincoln et al., 2005). In order to determine whether the effect of rs41295051 was independent of the main susceptibility locus, *HLA-DRB1*1501*, a stratified analysis was performed. Differential transmission of rs41295051 remained significant after correction in *HLA-DRB1*1501* negative MS affected offspring (Table 7). The effect of rs41295051 was no longer present in *HLA-DRB1*1501* positive offspring however (Table 8).

**Table 7 Family based test of genetic association of SNPs found within VDR binding sites and within or near MS associated genes stratified according to *HLA-DRB1*1501* in MS affected offspring (I).**

| **Transmissions to *DRB1*15* negative offspring** | | | | | | | |
|---|---|---|---|---|---|---|---|
| **SNP** | A1 | A2 | T | U | OR | CHISQ | P |
| **rs56383778** | C | G | 2 | 1 | 2.000 | 0.333 | 0.564 |
| **rs59608162** | - | - | 1 | 2 | 0.500 | 0.333 | 0.564 |
| **rs2243115** | G | T | 56 | 63 | 0.889 | 0.412 | 0.521 |
| **rs41295051** | A | G | 9 | 27 | 0.333 | 9.000 | 0.003 |
| **rs41295053** | G | A | 0 | 0 | NA | NA | NA |
| **rs34080929** | - | - | 0 | 0 | NA | NA | NA |
| **rs10905718** | G | A | 142 | 128 | 1.109 | 0.726 | 0.394 |
| **rs10747783** | - | C | 0 | 0 | NA | NA | NA |
| **rs2071170** | C | G | 6 | 12 | 0.500 | 2.000 | 0.157 |
| **rs3087456** | G | A | 115 | 111 | 1.036 | 0.071 | 0.790 |
| **rs6498139** | A | G | 40 | 45 | 0.889 | 0.294 | 0.588 |
| **rs4781028** | - | C | 0 | 0 | NA | NA | NA |
| **rs35494353** | - | - | 0 | 0 | NA | NA | NA |
| **rs428530** | G | A | 13 | 28 | 0.464 | 5.488 | 0.019 |
| **rs305083** | G | A | 95 | 103 | 0.922 | 0.323 | 0.570 |
| **rs182511** | G | C | 13 | 20 | 0.650 | 1.485 | 0.223 |
| **rs17081891** | - | T | 0 | 0 | NA | NA | NA |
| **rs17081900** | - | A | 0 | 0 | NA | NA | NA |
| **rs6566448** | - | T | 0 | 0 | NA | NA | NA |
| **rs17229961** | A | C | 27 | 16 | 1.688 | 2.814 | 0.093 |
| **rs280502** | T | G | 91 | 74 | 1.230 | 1.752 | 0.186 |

SNP = SNP identifier, A1 = minor allele code, A2 = major allele code, T = transmitted minor allele count, U = un-transmitted allele count, OR = TDT odds ratio, CHISQ = TDT chi-square statistic, P = TDT asymptotic p-value.

**Table 8 Family based test of genetic association of SNPs found within VDR binding sites and within or near MS associated genes stratified according *to HLA-DRB1*1501* in MS affected offspring (II).**

| **Transmissions to *DRB1*15* positive offspring** | | | | | | | |
|---|---|---|---|---|---|---|---|
| **SNP** | A1 | A2 | T | U | OR | CHISQ | P |
| **rs56383778** | C | G | 0 | 2 | 0.000 | 2.000 | 0.157 |
| **rs59608162** | - | - | 4 | 4 | 1.000 | 0.000 | 1.000 |
| **rs2243115** | G | T | 56 | 60 | 0.933 | 0.138 | 0.710 |
| **rs41295051** | A | G | 15 | 24 | 0.625 | 2.077 | 0.150 |
| **rs41295053** | G | A | 0 | 2 | 0.000 | 2.000 | 0.157 |
| **rs34080929** | - | - | 0 | 0 | NA | NA | NA |
| **rs10905718** | G | A | 150 | 149 | 1.007 | 0.003 | 0.954 |
| **rs10747783** | - | C | 0 | 0 | NA | NA | NA |
| **rs2071170** | C | G | 12 | 11 | 1.091 | 0.043 | 0.835 |
| **rs3087456** | G | A | 125 | 112 | 1.116 | 0.713 | 0.398 |
| **rs6498139** | A | G | 58 | 54 | 1.074 | 0.143 | 0.706 |
| **rs4781028** | - | C | 0 | 0 | NA | NA | NA |
| **rs35494353** | - | - | 0 | 0 | NA | NA | NA |
| **rs428530** | G | A | 23 | 21 | 1.095 | 0.091 | 0.763 |
| **rs305083** | G | A | 106 | 111 | 0.955 | 0.115 | 0.734 |
| **rs182511** | G | C | 17 | 16 | 1.062 | 0.030 | 0.862 |
| **rs17081891** | - | T | 0 | 0 | NA | NA | NA |
| **rs17081900** | - | A | 0 | 0 | NA | NA | NA |
| **rs6566448** | - | T | 0 | 0 | NA | NA | NA |
| **rs17229961** | A | C | 13 | 23 | 0.565 | 2.778 | 0.096 |
| **rs280502** | T | G | 83 | 98 | 0.847 | 1.243 | 0.265 |

SNP = SNP identifier, A1 = minor allele code, A2 = major allele code, T = transmitted minor allele count, U = un-transmitted allele count, OR = TDT odds ratio, CHISQ = TDT chi-square statistic, P = TDT asymptotic p-value.

***Predicted functionality of rs41295051***

The variant rs41295051 showed a protective effect in *HLA-DRB1*1501* negative affected offspring in this cohort. rs41295051 was previously located within a VDR peak after calcitriol stimulation in lymphoblastoid cells and ENCODE data shows that it lies within an open chromatin region marked by several histone modifications in multiple cells types (Boyle et al., 2012). VDR motifs are not predicted to occur over rs41295051 however (JASPAR, minimum relative threshold 70%). Its nearest gene, *IL2RA*, has its transcriptional start site 6877 bp downstream. IL2RA, an interleukin receptor, has previously been implicated in MS susceptibility by GWAS and is a target of MS treatment (De Jager et al., 2009) (Toubi et al., 2012). Independent replication and functional investigation of rs41295051 will be required in future studies.

# Discussion

The identification of genetic variation that disrupts transcription factor binding is important in efforts to define regulatory variants and fine map causal variants underlying GWAS. In these experiments I resolved genetic variation that is present within VDR binding intervals in human B cells and show two examples of disease associated SNPs, rs7859805 and rs1154154, lying within VDR motifs. Additionally, I show that rs41295051, a SNP lying within a VDR binding interval but not within a canonical motif, is significantly under-transmitted, conferring protection, in HLA-DRB1*15 negative affected offspring of MS families.

### AUTOIMMUNE DISEASE ASSOCIATED VARIANTS WITHIN MOTIFS DISRUPT VDR BINDING

For the variants rs7859805 and rs1154154, in near-perfect linkage disequilibrium with GWAS SNPs associated to CD and RA, and narcolepsy, respectively, I found that multiple motifs are present in their vicinity, making it more likely that genes nearby are under vitamin D regulation. Bioinformatic analysis predicted that the polymorphisms would cause changes in VDR binding affinity. I thus performed *in-vitro* experiments and show qualitative differences in binding due to these allelic variants. CD, RA and narcolepsy are diseases where vitamin D deficiency may be a contributing factor (Bronnum-Hansen et al., 2004; Lleo et al., 2008; McCombe et al., 2009; Ramagopalan et al., 2010). It is postulated that disruption of transcription factor binding could lead to altered gene expression and contribute to disease mechanisms (Knight et al., 1999). The electromobility shift assays (EMSA) performed here

support VDR binding disruption due to disease-associated variants in autoimmune conditions.

The ENCODE project has provided a wealth of data in numerous cell types using functional genomic methods and allows better interpretation of other experiments. The SNPs that I present in this study appear to be highly relevant as they are present in areas with features of regulatory regions in several cell types and disrupt motifs other than VDR. The SNPs rs7859805 (associated to CD and RA) and rs1154154 (associated to narcolepsy) are present in binding sites for multiple transcription factors and specific histone modifications. Moreover, rs7859805 is an eQTL for *PHF19*, a component of the polycomb repressive complex 2 that binds H3K36me3, a mark of active chromatin, and may be relevant in gene silencing during lineage specification (Brien et al., 2012). *PHF19* has been implicated in autoimmune disease and cancer by genome-wide association and functional studies (Brien et al., 2012). It is not known whether *PHF19* or other nearby genes to rs7859805 are regulated by vitamin D however. Less evidence of functionality is available for rs1154154 although it lies within a regulatory region where several other transcription factors such as SPI1 (PU.1), SRF and BCLAF1 bind.

#### GENETIC VARIANTS OUTSIDE OF MOTIFS MAY BE ASSOCIATED TO DISEASE

In the family based genetic association screen I tested variants found within VDR binding sites for transmission distortion in MS affected individuals. This analysis complements the findings of VDR binding disruption by showing that other variation

within binding intervals, not predicted to be in motifs, is also disease relevant and is likely to have functional impact.

Family based studies are a powerful method of assessing genetic disease association and can account for population structure (Spielman et al., 1996). The Canadian cohort is one of the largest in the world and has been on-going for more than 40 years. It has allowed epidemiological and genetic observations that have shed important insight into the aetiology and natural history of MS (Ebers et al., 1996; Sadovnick et al., 1996; Scalfari et al., 2010). This screen showed a significant association for rs41295051 in *HLA-DRB1*1501* negative offspring affected with MS. The SNP rs41295051 is found less than 7000 bp upstream of the MS implicated gene *IL2RA* but is not in high linkage disequilibrium with previously associated variants (rs12722489, rs2104286, rs706778 as described by Alloza et al. (2012)).

*IL2RA* has been implicated in MS susceptibility by genome-wide association and family based association studies (ANZgene, 2009; De Jager et al., 2009) (D'Netto et al., 2009). Here I find that the minor allele of a novel SNP found within a VDR binding site is protective against MS susceptibility in *HLA-DRB1*1501* negative affected offspring. The *IL2RA* gene codes the interleukin 2 receptor alpha and together with the IL2RB and IL2RG chains constitutes the IL2 receptor (IL2R). Mutations in *IL2RA* can cause interleukin 2 receptor alpha deficiency, a primary immune disorder characterised by abnormal B cell proliferation, normal B-cell development and decreased peripheral T cell numbers (OMIM #606367). IL2 signalling is central to the activation and regulation of the immune system and

increased expression of IL2RA in naive T cells, modulated by genetic variation, is known to increase the risk for MS (Gold et al., 2013). Laquinimod, a drug under current study for MS treatment, modulates B cells by increasing IL2RA in both B and T cells (Toubi et al., 2012). A phase II dose finding trial recently showed that daclizumab, a humanised monoclonal antibody that modulates interleukin-2 signalling by blocking IL2RA, has significant beneficial effects on MS activity as measured by relapse free periods and the annualised relapse rate (Gold et al., 2013). Calcitriol ($1,25(OH)_2D_3$) treatment increases IL2 expression, which in turn increases the proliferation of immune regulating $CD4^+CD25^+$ T cells (Gorman et al., 2010). Altered numbers and function of $CD4^+CD25^{high}$ cells and their modulation by IFNB treatment have been implicated in MS patient pathology (Sellebjerg et al., 2012). The effect of calcitriol on *IL2RA* expression has not been shown however.

**LIMITATIONS**

There are important limitations in this study including relevance of *in vitro* assays, power and information on allele frequencies. Several variants were absent or monomorphic in our MS cohort. Although closely related Caucasian populations have been described, more detailed knowledge of the genetic architecture will allow fewer assumptions (Abecasis et al., 2012). Importantly, replication in an independent cohort is necessary before further conclusions can be drawn.

Functional follow-up of allele specific binding represents a challenge. *In-vitro* methods for gene expression quantitation can be performed using cell lines and transfection assays. More recently, expression quantitative trait locus (eQTL) analysis

has emerged as a powerful tool to understand the genetics of gene expression. Collections of several tens to hundreds of individuals provide sufficient power to investigate cis acting eQTLs and can allow prioritisation of disease relevant loci for functional characterisation (Cookson et al., 2009; Majewski et al., 2011).

Methods to quantify differential binding are generally lacking and EMSA studies, despite its qualitative nature, represent the current standard. Novel methods interrogating allele specific binding are emerging and may allow more precise identification of relevant polymorphisms based on cohorts of previously genotyped individuals. Projects such as Hapmap and 1000 Genomes now allow these experiments to be carried out on genotyped and genome sequenced cell lines and the problems of sample collection, genotyping and sufficient quantities of cells are largely solved in this scenario.

Analysis of specific loci by quantification of chromatin immunoprecipitated (ChIP) DNA of individuals that carry different alleles at loci of interest may permit relative quantification of allele-specific binding (Knight et al., 2004; Knight et al., 2003). Genome-wide methods such as RNA-seq and ChIP-seq, with sufficient coverage and adequate mapping, may also allow unbiased relative quantification of allele-specific transcription (Lappalainen et al., 2013). More recent applications, such as ChIP followed by exonuclease digestion and next generation sequencing (ChIP-exo) increase the resolution of DNA-protein binding events to near single nucleotide level and may become the method of choice in the near future (Rhee et al., 2011). Statistical methods are only beginning to emerge for ASB analysis of genome-wide

data based on ChIP however and more robust methods will be required (Ji et al., 2013; Rozowsky et al., 2011).

## CONCLUSIONS

The molecular mechanisms underlying the effects of vitamin D in immune and inflammatory diseases are only beginning to be understood. Genetic variation appears to play an important role in this respect, both in regulating vitamin D and in modifying its effects at disease associated loci. Here I find that genome-wide binding of VDR occurs near MS and other disease associated variants more than expected by chance; that a polymorphism present in a VDR binding site significantly modulates susceptibility in a subset of MS patients; and, that autoimmune disease associated SNPs disrupt VDR binding at two loci.

Both rs1154154 and rs7859805 are in strong LD with disease-associated polymorphisms and cause VDR allele specific binding together with evidence of regulatory significance. Although I did not find variants associated to MS that potentially disrupt VDR motifs, this work uses current knowledge and resources to exemplify how genetic variation may be important in understanding the effects of vitamin D in immune diseases.

These results also serve to illustrate how family based tests of association can be used to analyse genetic variants that implicate biologically and disease relevant genes. In this family screen I identify rs41295051 as a novel genetic variant associated to MS susceptibility and VDR biology. Replication in an independent cohort is required however and further functional studies are needed to elucidate the mechanisms behind

VDR binding, genetic variation and IL2RA function in MS pathology if results are confirmed. The results provide information on specific loci that can be further investigated in disease cohorts and which may reveal important disease mechanisms. Future studies will also have to show the functional relevance in patient cohorts for rs7859805 and rs1154154 in CD, RA and narcolepsy.

Although each experimental method has important limitations, these studies serve to highlight the complexity of vitamin D physiology and the impact that disease associated genetic variation has in shaping its effects. These results may illustrate the importance of understanding transcription factor biology and the hypothesis that variation within and around motifs is disease relevant. They provide an approach to identify genetic variation that carries functional consequences and specifically show that genetic variation has an impact on the regulatory effects of vitamin D in immunity. In the following chapters I describe data derived from two cohorts that complement the laboratory based methods used here and further implicate vitamin D in immunity and inflammation.

***Molecular characterisation of genetic regulatory variants and vitamin D receptor binding in autoimmune disease:***

- rs1154154, a SNP near *TCRA* and associated to narcolepsy by GWAS, disrupts VDR binding;
- rs7859805, a SNP near *TRAF1* and associated to coeliac disease and rheumatoid arthritis by GWAS, also disrupts VDR binding;
- rs41295051, a SNP present in a VDR binding site near *IL2RA*, protects against MS susceptibility in HLA-DRB1*1501 negative individuals.

**Figure 18 Specific conclusions from results chapter three.**

# CHAPTER 4: INTEGRATIVE ANALYSIS OF VITAMIN D IN HEALTHY INDIVIDUALS

## Introduction

The following work represents a parallel approach to chapter three that allows studying the functional effects of vitamin D and its relationship to genotype and transcriptional regulation. Here I focus on understanding the associations of genetic variants, gene expression and plasma vitamin D by first looking within a cell specific context in healthy individuals and then analysing vitamin D in a cohort of patients with severe sepsis, an inflammatory condition with high mortality. The latter provides a translational focus to study the correlation of vitamin D, gene expression and phenotype. In these observational studies I use *in-vivo* methods to complement the molecular analyses carried out in the previous section. I provide molecular correlations to epidemiological observations using an integrative analysis within the context of immunity and inflammation. This work leverages hypothesis driven research with unbiased genome-wide tools to show that vitamin D is an important mediator of immunity in health and in disease.

Few studies have addressed the underlying pathophysiological mechanisms of GWAS signals and many associated genetic variants are thought to be in regulatory regions that could modulate transcription (Hindorff et al., 2009). GWAS of vitamin D have been published indicating that genetic variation is associated with circulating levels of $25(OH)D_3$ and that genotypic combinations increase risk of low levels (Ahn et al., 2010; Wang et al., 2010).

In a first instance I was interested in studying the effects of genetic variation on gene expression and plasma vitamin D levels in a cohort of healthy Caucasian volunteers for whom whole genome genotyping, $CD14^+$ monocyte and $CD19^+$ B cell transcriptomes and plasma were available. Other members of the J Knight Lab carried out an expression quantitative trait study (Fairfax et al., 2012) in these individuals and in this chapter I present the investigation of vitamin D in the same cohort with correlations to genotype and transcriptome.

Given the current debate regarding the level of $25(OH)D_3$ necessary for health maintenance, for this study I defined the normal range of vitamin D in plasma to be 50–150 nmol/L with deficiency thresholds at <25 (risk of vitamin D deficiency rickets increases) and <50 nmol/L (production of $1,25(OH)_2D_3$ and PTH measured in plasma becomes impaired) and insufficiency between 51-75 nmol/L (>75 nmol/L reduces the risk of falls, fractures and colorectal cancer) (Ross et al., 2011). Upper thresholds of normality are not clearly defined although levels in healthy individuals derived from sun exposure can reach 235 nmol/L (Ross et al., 2011).

In this chapter I explore vitamin D deficiency GWAS signals, replicate previous findings and correlate cell specific gene expression with plasma vitamin D. In the following chapter, I apply a similar approach in a cohort of patients with severe sepsis. As with healthy individuals, I also identify a correlation between gene expression and plasma vitamin D and additionally show several molecular associations with vitamin D deficiency and increased mortality that provide support to the clinical and epidemiological observations.

Although these studies have important limitations they provide insight into vitamin D actions and a functional molecular link that informs candidate genes involved in sepsis pathophysiology and outcome.

**Integrative analysis of vitamin D in healthy individuals:**

- What is the distribution of plasma levels of vitamin D in healthy individuals?
- Can I replicate GWAS findings of variants associated to vitamin D deficiency using a candidate approach in a cohort of healthy individuals?
- Is there a relationship between gene expression and plasma levels of 25-hydroxyvitamin D in healthy volunteers? Is this relationship cell specific?
- Does pathway analysis reveal informative networks for disease associated research?

**Figure 19 Questions addressed in chapter four.**

# Results

## HEALTHY VOLUNTEERS COHORT CHARACTERISTICS AND PLASMA VITAMIN D LEVELS

Healthy Caucasian individuals had been recruited from the Oxfordshire area following informed consent. Whole blood was extracted and cells were positively separated using antibody coated magnetic beads. Plasma, $CD14^+$ monocyte RNA, $CD19^+$ B cell RNA and DNA were available for 257 volunteers after sample processing, quality control and outlier detection from an initial group of 288 (see Chapter 2: Materials and Methods). In the final cohort, there were 112 (44%) male and 145 (56%) female volunteers aged 18 to 62 years old with an average of 33 years. Blood samples were taken between March and July 2010. Whole genome SNP array genotyping was performed on Illumina OmniExpress and cell transcriptomes on Illumina BeadArray HT-12v4. Albumin, C-reactive protein (CRP) and $25(OH)D_2$ and $D_3$ plasma measurements were carried out in a clinically approved laboratory. Vitamin D metabolites were measured using the current gold standard tandem mass spectrometry liquid chromatography. CRP in plasma was used to identify individuals with underlying inflammatory processes with non-detectable levels in all patients (see Chapter 2: Materials and Methods).

I performed an exploratory analysis of plasma $25(OH)_2D_3$ and found that the healthy volunteer cohort had a median of 58 nmol/L (inter-quartile range 43-74 nmol/L; mean 60 nmol/L) with one individual above 150 nmol/L (156 nmol/L) (Figure 20).

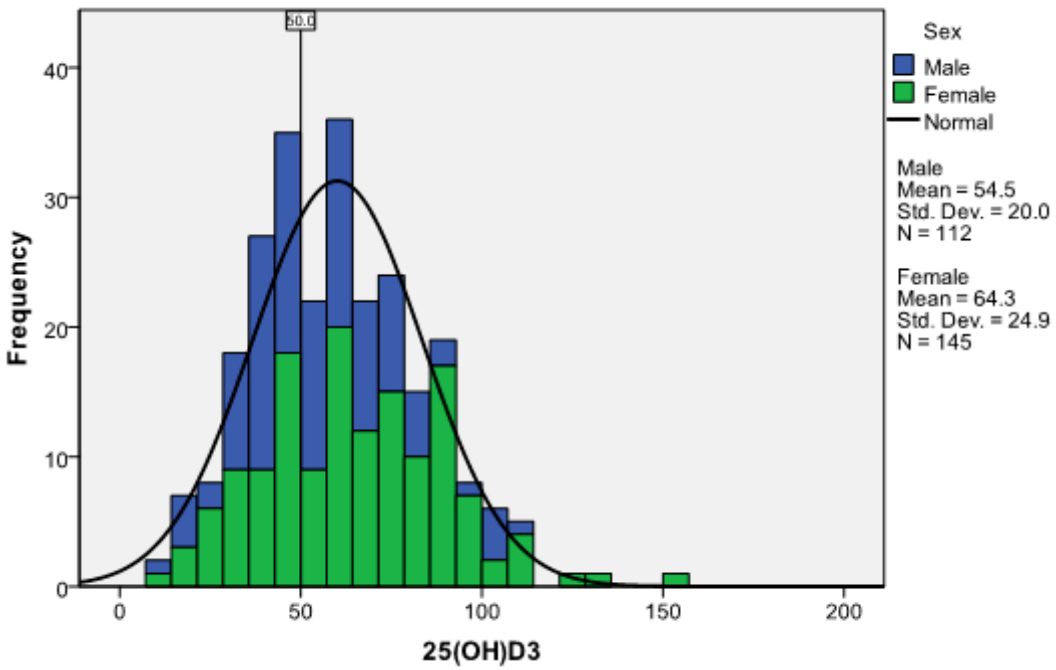



**Figure 20 Frequency distribution of plasma levels of 25(OH)$D_3$ in 257 healthy volunteers.**

Vitamin D levels in blood were measured in healthy individuals. The histogram shows male and female distributions. The bar indicates 50 nmol/L, the threshold for vitamin D deficiency. Mass spectrometry was used to measure 25(OH)$D_3$ with a lower detection limit of 10 nmol/L.

Plasma levels according to different thresholds of 25(OH)$D_3$ are shown in Table 9. Taking March and April as the end of winter and July as summer (closest measurements in this cohort to accepted definitions), I found that 92% and 60% had 75 nmol/L or below respectively. These values are comparable to large Caucasian and British cohorts previously studied (Hypponen et al., 2007; Wang et al., 2010). Levels are expected to increase over spring and summer and may show gender differences (Hypponen et al., 2007).

**Table 9 Plasma levels of 25(OH)$D_3$ in healthy individuals.**

| Plasma levels of 25(OH)$D_3$ in healthy individuals | | | | | | |
|---|---|---|---|---|---|---|
| **All samples** | | | March and April | | July | |
| **nmol/L** | n | % | n | % | n | % |
| **<=25** | 11 | 4.3 | 11 | 16.7 | 1 | 1.4 |
| **<=50** | 97 | 37.7 | 45 | 68.2 | 14 | 20.0 |
| **51to75** | 98 | 38.1 | 16 | 24.2 | 28 | 40.0 |
| **>75** | 62 | 24.1 | 5 | 7.6 | 28 | 40.0 |
| **Total** | 257 | 100 | 66 | 100 | 70 | 100 |

Values were grouped according to deficiency (<25 or <50 nmol/L), insufficiency (51 to 75 nmol/L) and sufficiency (>75 nmol/L).

When comparing the levels of 25(OH)$D_3$ according to the month of sampling I observed, as expected, significant differences between individuals sampled in April (n=66) (only 8 individuals were recruited in March) compared to July (n=70) (median 43.7 nmol/L v. 69.7 nmol/L respectively, p-value <0.001). I compared the distribution according to gender and found that females had significantly higher values (p-value=0.001; exact two tailed p values, Mann Whitney U tests) (Figure 9 and appendix Table 27). This data is consistent with previous evidence suggesting that levels are expected to increase over spring and summer and may show gender differences (Hypponen et al., 2007).

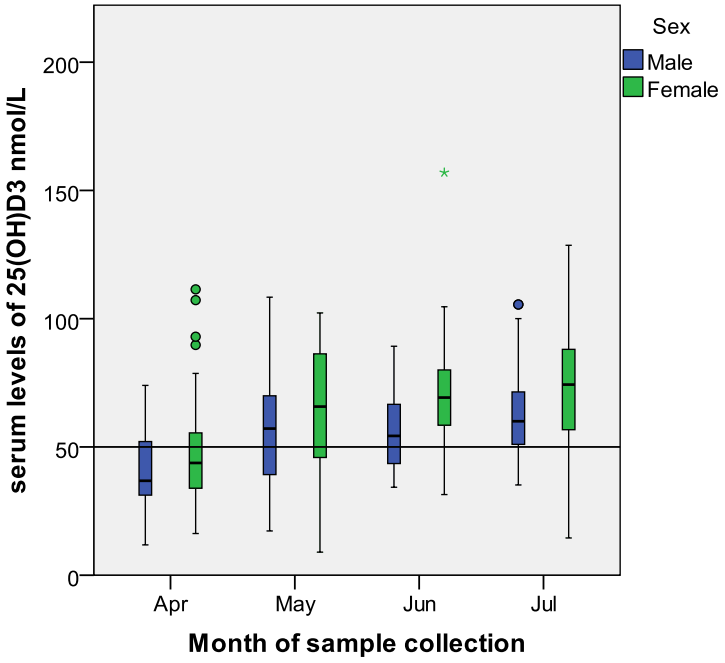



**Figure 21 Boxplots of 25(OH)$D_3$ plasma levels according to gender and month of sampling.**

Median and interquartile (IQR) ranges shown; the ends of the whiskers represent the lowest/highest data point within 1.5 IQR of the lower/upper quartile). Samples were grouped according to month of sampling and gender. March (n=8) and April are shown together. Asterisk/circle=outliers. See Table 28 in appendix for details.

### QUANTITATIVE TRAIT ANALYSIS OF PLASMA VITAMIN D

After exploring the distribution of vitamin D levels I looked at the possible association of vitamin D and genetic variation with the aim of replicating findings from previous GWAS (Ahn et al., 2010; Wang et al., 2010). Although whole genome genotyping had already been performed as part of a separate eQTL study of global gene expression at the transcript level (Fairfax et al. 2012), a candidate variant

approach was used to conserve power. The frequency of three specific variants previously associated by GWAS to vitamin D levels were analysed (Wang et al., 2010). These variants are in close proximity to genes known to be essential in vitamin D metabolism including *CYP2R1*, encoding vitamin D 25-hydroxylase (rs2060793); *GC*, encoding the vitamin D binding protein (rs2282679); and *DHCR7*, encoding 7-dehydrocholesterol reductase (rs3829251) (Ahn et al., 2010; Wang et al., 2010). Two of these SNPs, rs2060793 and rs2282679, were found to show statistically significant association with plasma vitamin D levels (raw p-values <0.05) (Table 10). A third, rs3829251, was underpowered to detect differences in homozygotes for the risk genotype (A/A). All SNPs showed the reported trend in modulating plasma vitamin D levels.

**Table 10 Single nucleotide polymorphisms determine the levels of plasma vitamin D in healthy Caucasians.**

| SNP and risk allele | rs2060793-A | | | rs2282679-G | | | rs3829251-A | | |
|---|---|---|---|---|---|---|---|---|---|
| **Associated gene*** | CYP2R1 | | | GC | | | NADSYN1,DHCR7 | | |
| **P-value in this study** | 0.003388 | | | 0.008538 | | | 0.3903 | | |
| **Genotype** | A/A | A/G | G/G | G/G | G/T | T/T | A/A | A/G | G/G |
| **Number of individuals** | 43 | 110 | 100 | 24 | 115 | 114 | 5 | 58 | 190 |
| **Frequency** | 0.17 | 0.44 | 0.4 | 0.1 | 0.46 | 0.45 | 0.02 | 0.23 | 0.75 |
| **25(OH)$D_3$ nmol/L mean** | 22.3 | 20 | 18.5 | 17.9 | 18.9 | 21.1 | 17.5 | 19.4 | 20 |
| **SD** | 6.7 | 7.4 | 7.1 | 7.3 | 6.7 | 7.6 | 4 | 8.1 | 7.1 |
| **Reported tendency** | increases | | | decreases | | | decreases | | |

The cohort comprised 257 volunteers. PLINK was used to perform the quantitative trait analysis between genotype and plasma 25(OH)D values (Purcell et al., 2007). *By proximity and known gene function.

### VITAMIN D IS ASSOCIATED TO CELL SPECIFIC TRANSCRIPTOMES

In order to investigate possible mechanisms of gene regulation by vitamin D, plasma levels of 25(OH)$D_3$ were correlated to cell type specific genome-wide gene expression levels that had been assayed previously (Table 11 and appendix Table 29). Whole genome gene expression was quantified using Illumina HumanHT-12 v4 BeadChip gene expression arrays and plasma levels of vitamin D were measured from the same volunteer samples. Background subtraction, batch and array effect removal

were carried out using R packages; data was normalised using the robust spline method and standard quality control was performed. A total of 29,022 probes remained for further analysis for 286 B-cell and 287 monocyte samples (Fairfax et al., 2012) (see Chapter 2: Materials and Methods for full details). For the transcriptome-plasma vitamin D analysis Spearman's rank correlation in the statistical package R was used. Using a false discovery rate (FDR) of <0.05, I found expression of only 12 genes correlated with plasma vitamin D in B cells versus 587 for monocytes. The average correlation (rho) at an FDR of <0.05 was moderate with 0.28 for B cells and 0.23 for monocytes (absolute values) and comparable to studies with a similar approach (Ghazalpour et al., 2011).

**Table 11 Results of B cell and monocyte transcriptome correlations with 25(OH)$D_3$ plasma levels from healthy volunteers.**

| (A) B cells, all correlated genes | | | (B) Monocytes, top 15 genes | | |
|---|---|---|---|---|---|
| **Gene symbol** | **p-value** | **R (rho)** | **Gene symbol** | **p-value** | **R (rho)** |
| DCTN6 | 2.61E-06 | -0.29 | ZNF259 | 4.29E-07 | -0.31 |
| SLC35E3 | 3.34E-06 | -0.28 | C1ORF57 | 6.77E-07 | -0.3 |
| IGFL1 | 5.71E-06 | 0.28 | CCDC86 | 8.18E-07 | -0.3 |
| LOC644086 | 6.48E-06 | 0.28 | TAF5L | 1.12E-06 | -0.3 |
| LOC100129158 | 7.10E-06 | -0.28 | SEH1L | 1.41E-06 | -0.29 |
| NUP50 | 7.40E-06 | -0.27 | SCYL1 | 1.45E-06 | -0.29 |
| TXNRD3IT1 | 7.60E-06 | -0.27 | RNU1-5 | 1.54E-06 | 0.29 |
| IL27RA | 9.01E-06 | -0.27 | TMEM126A | 1.83E-06 | -0.29 |
| ERLIN1 | 9.09E-06 | -0.27 | ALG8 | 2.80E-06 | -0.29 |
| LOC647650 | 9.88E-06 | 0.27 | MRPL12 | 3.15E-06 | -0.28 |
| TLR7 | 1.04E-05 | -0.27 | HS.36053 | 3.85E-06 | 0.28 |
| WDR70 | 1.17E-05 | -0.27 | THYN1 | 4.74E-06 | -0.28 |

**(C) Genes of interest, monocyte-25(OH)D correlated genes**

| **Gene symbol** | **p-value** | **R (rho)** |
|---|---|---|
| HIST1H1C | 2.76E-05 | 0.257 |
| JUN | 1.38E-05 | 0.267 |
| NFKBIB | 6.26E-04 | -0.211 |
| TGFBI | 1.69E-04 | -0.232 |
| IL27RA | 3.25E-04 | -0.222 |

Spearman's correlation for expression of 29,022 probes with plasma vitamin D is shown for (A) B cells (all correlated genes) and monocytes ((B) top and (C) genes of interest from a total of 587 correlations) at a p-value <0.05 and FDR < 0.05. See appendix Table 29 for the full list of genes for monocyte correlations.

### PATHWAY ANALYSIS ASSOCIATES VITAMIN D WITH TLR, TNF AND MAPK SIGNALLING

To understand further the relationship of correlated genes with vitamin D levels, I performed pathway analysis for each cell type (IPA v8.8 Ingenuity Systems) (Table 12). No significant results were found for B cells at an FDR of <0.05 (n=12) so I used a higher threshold (FDR <0.10, p-value <0.05, n=1510) for this cell type. Both cell types showed significant enrichment for genes present in the lipopolysaccharide (LPS) stimulated mitogen activated protein kinase (MAPK) signalling pathway (IPA, corrected p-value <0.05). MAPK signalling is essential in inflammatory responses and strongly implicated in sepsis (Risco et al., 2012) and MS (Kleinewietfeld et al., 2013). Other pathways of importance such as Toll like receptor signalling in B cells and TNF receptor I (TNFRSF1A) signalling in monocytes were also highlighted as significant (Table 12).

**Table 12 Significant canonical pathways of transcriptome-25(OH)$D_3$ correlated genes.**

| B cell transcriptome-25(OH)$D_3$ (FDR <0.1) | | Monocyte transcriptome-25(OH)$D_3$ (FDR <0.05) | |
|---|---|---|---|
| **Pathway** | **P value** | **Pathway** | **P value** |
| Toll like receptor signalling | 1.19E-07 | RANK signalling in osteoclasts | 6.6E-06 |
| Production of nitric oxide and reactive oxygen species in macrophages | 2.22E-07 | CD27 signalling in lymphocytes | 8.54E-06 |
| LPS-stimulated MAPK signalling pathway | 1.32E-06 | TNFR1 signalling | 3.58E-05 |
| Apoptosis signalling | 2.37E-06 | Rac signalling | 1.38E04 |
| B cell receptor signalling | 6.98E06 | LPS stimulated MAPK signalling | 1.6E-04 |

I performed pathway analysis using Ingenuity Systems (IPA, build version 8.8) with the following settings: reference set whole genome, direct and indirect relationships, endogenous chemicals included and considering only molecules and relationships experimentally observed in humans.

# Discussion

In this study plasma 25(OH)$D_3$ was measured in a cohort of healthy British Caucasians and correlated to candidate genotypes and cell-specific transcriptomes. I found that a significant proportion of individuals had deficient or insufficient levels of plasma vitamin D. These levels were comparable to a large study in healthy white British individuals which observed proportions of 87% and 61% for end of winter and end of summer at the same cut-offs (Hypponen et al., 2007). I observed differences in gender which have also been reported before in healthy populations and are thought to depend on season, BMI, dietary and supplementation factors amongst others (Hypponen et al., 2007).

I explored the contribution of genetic variation in determining plasma vitamin D levels. Previous findings from GWAS studies were replicated and implicate genes known to be essential in vitamin D metabolism, *CYP2R1*, encoding vitamin D 25-hydroxylase, and *GC*, encoding the vitamin D binding protein (Ahn et al., 2010; Wang et al., 2010). Scarce evidence exists as to how these variants may exert their regulation over plasma vitamin D however. The SNPs are not known to act as expression quantitative trait loci although they lie within regulatory regions in multiple cells types (Boyle et al., 2012; Dunham et al., 2012). These SNPs, or variants in high linkage disequilibrium with them (specifically rs1155563, rs2276360 and rs1037379), are predicted to disrupt binding of the transcription factors POU2F2 (OCT2), CEBPB and STAT (RegulomeDB analysis, data not shown). Although little is known, these factors are associated with vitamin D and may be in part responsible

for determining levels of the implicated vitamin D metabolism genes (*CYP2R1, GC* and *DHCR7*) (Kakizawa et al., 1999; Liu et al., 1994). In monocytes, interferon-gamma and CD14/TLR4 induction of *CYP27B1*, the vitamin D activating enzyme, requires the JAK-STAT, NF-kappaB, and p38 MAPK pathways as well as binding of CEBPB (Stoffels et al., 2006).

Few studies have addressed the relationship between metabolites and gene expression. In this cohort of healthy volunteers I studied the correlation between cell specific gene expression, derived from $CD14^+$ monocytes and $CD19^+$ B cells, and plasma $25(OH)D_3$. This analysis is in agreement with a publication in mice showing that significant correlations are not necessarily strong between expressed genes and protein abundance (~r^2=+/-0.25) although the meaning of this is unclear (Ghazalpour et al., 2011).

Genes expressed in B cells that were correlated with plasma $25(OH)D_3$ at statistically significant levels included disease relevant molecules such as *TLR7, IL27RA* and *JUN* (*AP1*). Expression of *IL27RA* is known to be modulated by $1,25(OH)_2D_3$ in monocytes, a VDR binding interval is present near the transcriptional end site of *TLR7* in B cells and the expression of *JUN* family members is also affected by $1, 25(OH)_2D_3$ (Heikkinen et al., 2011; Ramagopalan et al., 2010; Wang et al., 2005). *TLR7,* encoding Toll like receptor 7, is a pattern recognition receptor essential in pathogen detection. TLR7 recognises single stranded RNA in endosomes, a common viral signature. It is known to be expressed in B cells and appears to play a role in B cell self-reactivity to RNA (Green et al., 2011). Indeed, TLR7 has been associated to

several autoimmune disorders, including MS where its deficiency in patients can be rescued with IFN-beta therapy (Giacomini et al., 2013; Rafalski et al., 2013). Using the same approach I show in the next chapter that leukocyte *TLR7* expression in patients with severe sepsis or septic shock is also significantly correlated to plasma levels of 25(OH)D and to those of the antimicrobial peptide cathelicidin.

Vitamin D - transcriptome correlations in monocytes were far more pronounced with over 500 associated genes at significant thresholds. Genes of importance in inflammatory responses and disease included *IL27RA*, also correlated to vitamin D in B cells, which encodes the alpha unit of the interleukin 27 receptor. IL27RA (also known as WSX1 and TCRR) is necessary for Th1 cell differentiation through its actions on STAT1 and T-bet (Takeda et al., 2003). IL27RA deficiency leads to impaired IL-12-dependent IFN-gamma production and altered Th1/Th2 T cell responses (Takeda et al., 2003). Aberrant Th1 T cells are implicated in autoimmunity and the protective effects of vitamin D are thought to occur through modulation of regulatory T lymphocytes and Th1/Th2 T cell balance (Antico et al., 2012; Baeke et al., 2010).

Other genes of known importance in immune regulation that were significantly correlated to vitamin D in monocytes included members of the nuclear factor kappa B (*NFKBIB)*, transforming growth factor (*TGFBI*), and interferon regulatory factor (*IRF2*) families. Molecules from these important families have been previously associated to regulation by vitamin D (Pike et al., 2010; Ramagopalan et al., 2010). In the next chapter I also show that interferon family members, including *IRF2,* and

members of the NFKB family, are also significantly correlated to plasma vitamin D in patients with severe sepsis.

Analysis of transcriptome-vitamin D correlations revealed canonical pathways that are relevant to disease and that could inform investigations in disease cohorts. The LPS-MAPK signalling pathway was highlighted in both cell types and is interesting given the association of vitamin D to severe sepsis. This is also in accordance with vitamin D playing an important role in immune function.

Fewer genes expressed in B cells were correlated with plasma $25(OH)D_3$ compared to monocytes. Differences in cell-specific transcriptomes may relate to the level of responsiveness of particular cell types and to specific regulatory mechanisms. Calcitriol induces monocyte maturation and promotes phagocytic and chemotactic macrophage responses while calcifediol (25(OH)D) upregulates *CYP27B1* in the presence of interferon gamma in monocytes (Hewison, 2011). In B cells calcitriol is known to inhibit proliferation, plasma-cell differentiation and immunoglobulin production (Baeke et al., 2010). Questions in cell specific effects driven by transcription factor function remain open and much further research is required.

These results provide evidence of how integrating diverse levels of biological information can give insight into vitamin D biology. It is known that genetic variation plays a role in gene expression and that gene expression and metabolite levels interplay. Future work, including for example systems level perturbation with identified cellular phenotypic endpoints, could help elucidate the exact mechanisms whereby regulation of and by vitamin D in relevant inflammatory and immune

contexts occurs. Recent work indicates that mouse models poorly reflect human inflammatory states (Seok et al., 2013). Integrated approaches may be of increased utility to identify relevant disease pathways, mechanisms, druggable targets and classification systems in complex conditions where direct observations in humans are necessary. In the following chapter I show how this method, coupled to risk factor identification, can provide potentially translatable insight in severe sepsis.

**Integrative analysis of vitamin D in healthy individuals:**

- Common genetic variation helps determine plasma vitamin D levels;
- Plasma vitamin D significantly correlates to cell specific gene expression in immune cells;
- Vitamin D correlated genes derived from cell specific transcriptomes highlight statistically significant pathways important in immune and inflammatory responses such as MAPK signalling.

**Figure 22 Specific conclusions from chapter four.**

# CHAPTER 5: CLINICAL AND MOLECULAR CORRELATES OF VITAMIN D AND SEPSIS SURVIVAL

## Introduction

Sepsis is defined as a dysregulated systemic inflammatory response to infection with complex dynamic pathophysiological mechanisms in which there is a continuum from sepsis to septic shock (Quesada et al., 1995). It has a high mortality rate and therapeutic options are currently limited. Recent research has highlighted the role of vitamin D as an important modulator of immunity with both anti-inflammatory and anti-microbial effects (Baeke et al., 2010; Edfeldt et al., 2010; Fabri et al., 2011; Hewison, 2011). Deficiency of vitamin D is common and has been linked to many diseases through epidemiological studies (Holick, 2007; Ramagopalan et al., 2010). Many hospitalised patients are vitamin D deficient and deficiency of this vitamin is severe among ICU patients (Lee et al., 2009; Thomas et al., 1998). Seasonal and geographic variation in sepsis incidence and mortality with increased rates in winter, not fully accounted for by changes in viral respiratory infections, and a possible latitude effect have been described (Danai et al., 2007).

Given the pleiotropic effects of vitamin D and its growing recognition as a modulator of innate and immune responses I hypothesized that vitamin D deficiency might be a contributing factor to all-cause mortality in patients with severe sepsis/septic shock. Previous studies investigating vitamin D levels in patients with sepsis have had small sample sizes, lacked molecular characterisation and/or long term follow-up. Here, I combine clinical observations with a functional genomics approach to understand

disease mechanisms in patients treated in the ICU (see Figure 24 for the analytical strategy). I also explore possible mechanisms through which vitamin D could influence biological pathways and outcome in these patients by relating global vitamin D levels to gene expression profiles.

In chapter one I provide detailed introductions to sepsis and vitamin D and in chapter two I describe in full the materials and methods used. In this section I present the results of this investigation and show that deficiency of vitamin D is severe in sepsis patients and correlated to increased risk of death. An integrated analysis of plasma vitamin D, polymorphonuclear cell counts, genome-wide gene expression and clinical and biochemical parameters revealed a patient cluster at increased risk of early death. This patient cluster was not evident as measured by disease severity (SOFA), prognostic scores (APACHE) or other routinely collected parameters.

Sepsis has suffered from many failed clinical trials. Integrated approaches with detailed phenotyping may aid in identifying disease mechanisms and in more accurately classifying patients for therapeutic and prognostic purposes.

**Clinical and molecular correlates of vitamin D and sepsis survival:**

- Are patients with severe sepsis or septic shock vitamin D deficient? If so, is this associated with survival? Can vitamin D levels help predict sepsis outcome?
- Can integrating gene expression, plasma vitamin D levels and relevant clinical and molecular information help identify sub-groups of patients?

**Figure 23 Questions addressed in chapter five.**

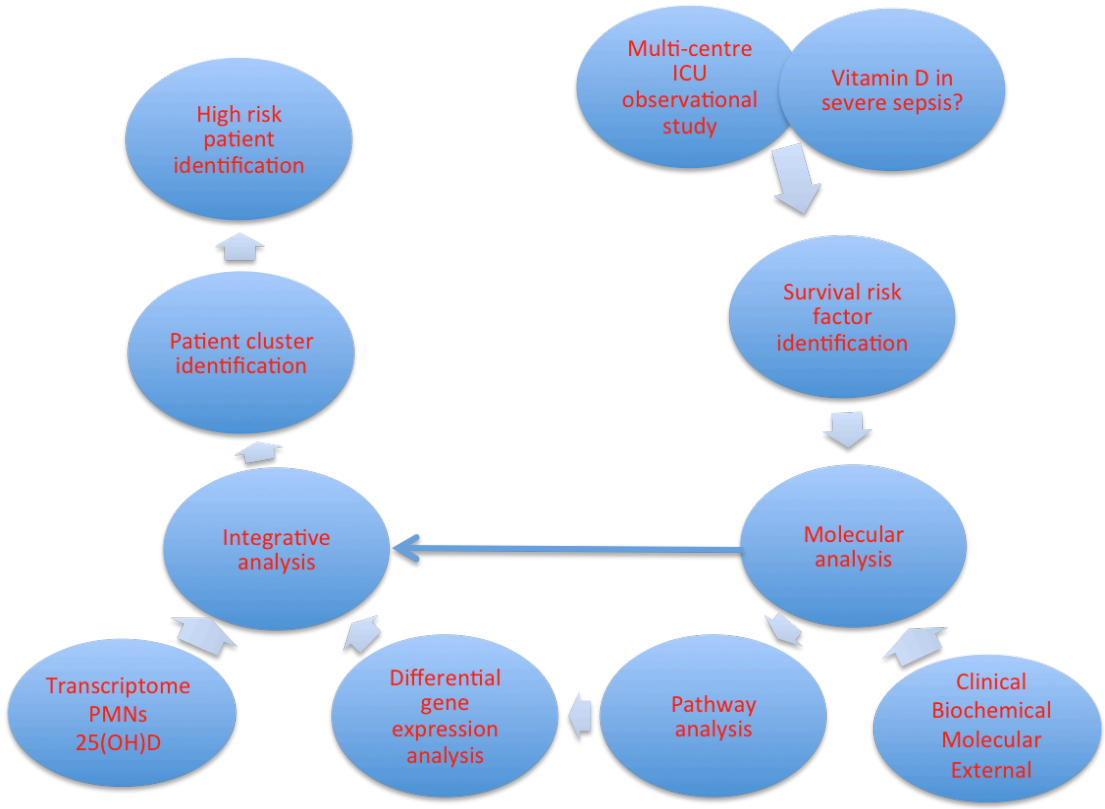



**Figure 24 Analysis workflow in the GAinS vitamin D study.**

Molecular analysis refers to clinical, biochemical and molecular (gene expression and peptide measurements not investigated routinely) collected during hospital, ICU stay and follow-up of the GAinS cohort as detailed in the main text and in chapter two. External refers to publicly available datasets used for intersection with GAinS generated data. Transcriptome refers to leukocyte genome-wide gene expression from GAinS patients. 25(OH)D was measured in plasma. PMNs = polymorphonuclear cells; ICU = intensive care unit.

# Results

## PARTICIPANTS AND DESCRIPTIVE DATA

This study formed part of the Genomic Advances in Sepsis (GAinS) project, a UK multi-centre ICU research collaboration (http://www.ukccg-gains.org/). Ethical approval was granted and patient recruitment was carried out from 2005 until 2011. Adults treated in the ICU and diagnosed with sepsis as defined by international guidelines (Annane et al., 2005) were eligible (see Appendix Table 19 for clinical criteria of sepsis). The study enrolled 144 consenting patients of whom one withdrew after consenting and was excluded from analyses and 2 were lost to follow-up at 6 months survival. Patients had severe sepsis due to either community-acquired pneumonia (CAP; n=77) or faecal peritonitis (FP; n=66). CAP and FP definitions, exclusion criteria and further methodological details are described in chapter two. Sepsis spectrum, mechanisms, definitions and epidemiology are described in chapter one. The Appendix additionally contains the clinical criteria and definition of the systemic inflammatory response syndrome and sepsis spectrum (Table 19), mechanisms involved in tissue damage from sepsis (Table 20) and major systems and organs affected (Table 21). Treating ICUs had subscribed to management according to Surviving Sepsis Campaign Guidelines (Dellinger et al., 2013).

Patients were diagnosed with severe sepsis (42%) or septic shock (58%, as determined by vasopressor use, see appendix Table 31). All patients from this cohort can be encapsulated by severe sepsis and I thus use this term from here on to refer to all cases for simplicity.

Patients had a median age of 64.5 years, 51% were male, 93% Caucasian and an overall mortality rate at hospital discharge of 24.3% and at 6 months of 29.8% (see appendix Table 30 and Table 31).

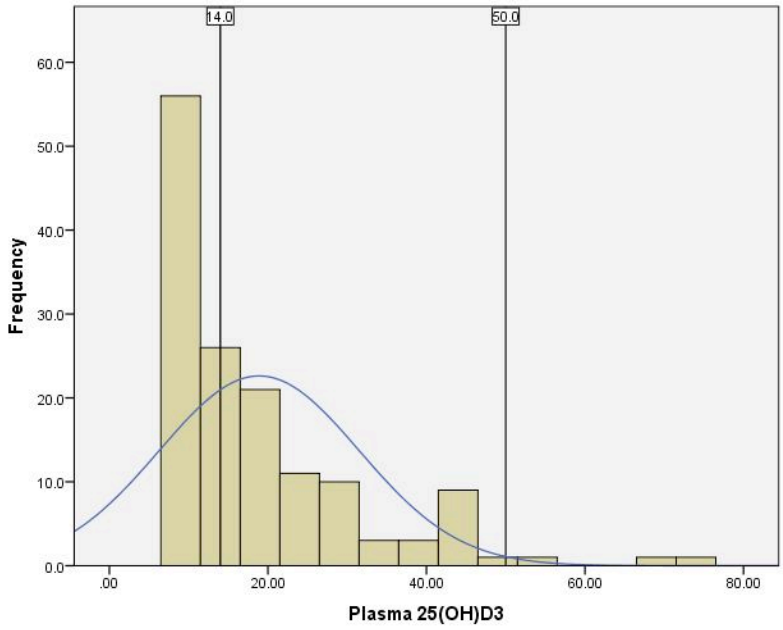



**Figure 25 Frequency distribution of plasma levels of 25(OH)D in patients with severe sepsis.**

Samples represent the first available from ICU admission (day 1, 3 or 5) and are 25(OH)$D_2$ plus 25(OH)$D_3$ (referred to as 25(OH)D) to reflect total vitamin D biological activity. Although largely equivalent, 25(OH)$D_2$ (dietary) and 25(OH)$D_3$ (sunlight exposure) have different sources and pharmacological properties and must be accounted for (Armas et al., 2004). The first bar indicates the median (14 nmol/L) in this cohort and the second bar 50 nmol/L, the current cut-off for vitamin D deficiency (Ross et al.,

2011). Mass spectrometry was used to measure 25($OHD_2$ and 25(OH)$D_3$ in plasma with a lower detection limit of 10 nmol/L.

## OUTCOME MEASURES AND MAIN RESULTS

### *Plasma vitamin D deficiency is severe, common and associated with death in patients with severe sepsis*

Analysis of plasma levels of 25(OH)D showed that 97% of patients were deficient (<50 nmol/L) and 77% severely deficient (<25 nmol/L) (Figure 25 and appendix Table 32). I performed univariate analysis of 25(OH)D using demographic and clinical characteristics considered to be important for vitamin D biology and sepsis prior to multivariate procedures.

I found significantly lower levels among patients who died compared to survivors both at hospital discharge (median 10 nmol/L vs. 16 nmol/L, P=0.011) and at 6 months after discharge (median 11 nmol/L vs. 16 nmol/L, P=0.003) but not at ICU discharge or 28 days (P=0.166 and P=0.121 respectively; Mann Whitney U tests, 2-tailed exact P values).

The half-life of vitamin D (as the 25(OH)D precursor) is long (~8 weeks (Jones, 2008)) but recent studies indicate that consumption may occur in acute states (Bang et al., 2011). Significance values of the association of low plasma 25(OH)D to increased death were not materially modified by the day of sampling (appendix Table 33 and Table 34). A subset of patients had plasma 25(OH)D measurements available for days 1, 3 and/or 5 of ICU admission. Vitamin D levels on all days were significant at 6 month mortality with higher 25(OH)D in survivors (P=0.016, 0.040, 0.021,

respectively; appendix Table 33 and Table 34) and showed similar results when testing proportions based on the median and day of sampling (Table 35 and Table 36).

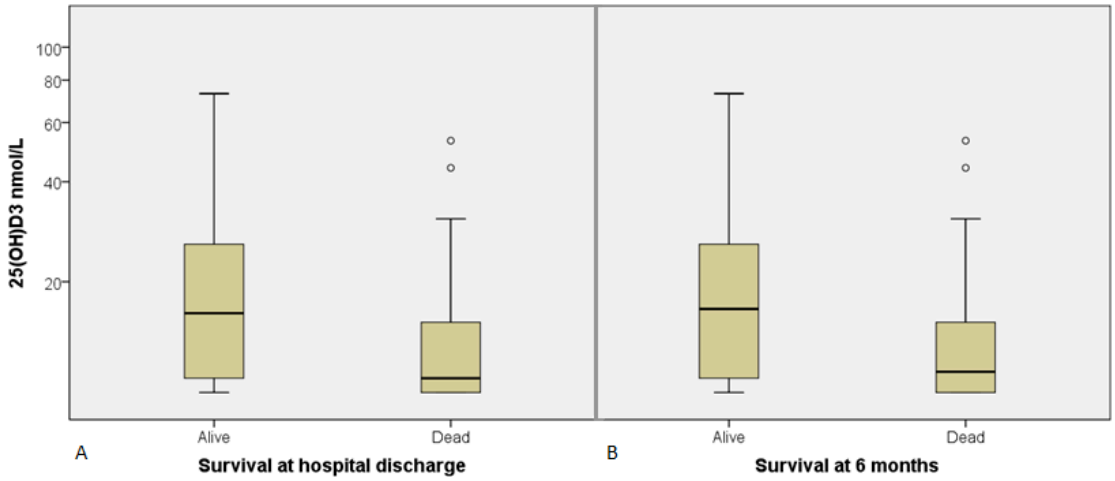



**Figure 26 Plasma levels of 25(OH)$D_3$ in patients with severe sepsis or septic shock according to survival at (A) hospital discharge, and (B) 6 months following ICU admission.**

Median and interquartile (IQR) ranges shown; the ends of the whiskers represent the lowest/highest data point within 1.5 IQR of the lower/upper quartile). Y-axis plotted on log scale base10, circle = outlier.

***Sub-group analysis***

Seasonal variation in sepsis incidence and mortality has been observed previously in large retrospective studies (Danai et al., 2007). Although power may be limited in this cohort I analysed patients by season of admission. Plasma levels of 25(OH)D of patients admitted to ICU during winter or spring were significantly lower (median 12 nmol/L) than those admitted during summer or autumn (median 17 nmol/L; P=0.014). The proportion of deaths according to season in this cohort was not significantly

different however (winter + spring vs. summer + autumn, Pearson chi square test > 0.05) (appendix Table 35). Most patients had severe vitamin D deficiency with many below the detection limit during winter and autumn. Comparisons according to gender (P=0.52), age (≤64 years vs. >64 years, P=0.56) and diagnosis (CAP vs. FP, P=0.076) were not significant (Mann Whitney U test, exact two tailed P values).

**Table 13 Distribution values of plasma 25(OH)D according to survival in patients with severe sepsis**

| **Plasma 25(OH)D (nmol/L)** | **Survival at ICU discharge** | | **Survival at 28 days** | | **Survival at hospital discharge** | | **Survival at 6 months** | |
|---|---|---|---|---|---|---|---|---|
| | Alive | Dead | Alive | Dead | Alive | Dead | Alive | Dead |
| Valid N | 120 | 23 | 113 | 28 | 108 | 35 | 99 | 42 |
| Median | 14.0 | 11.0 | 14.0 | 11.0 | 15.5 | 10.0 | 16.0 | 10.5 |
| Percentile 25 | 9.0 | 9.0 | 9.0 | 9.0 | 10.0 | 9.0 | 10.0 | 9.0 |
| Percentile 75 | 25.0 | 15.0 | 25.0 | 16.0 | 25.5 | 15.0 | 26.0 | 15.0 |

***Regression analysis to predict outcome from severe sepsis***

In order to assess whether other variables had a larger influence or confounded the univariate results on survival outcome I performed binary logistic regression. This analysis took into account known confounders for vitamin D such as season of admission, age and gender, as well hospital length of stay, comorbidity (Charlson index), organ failure (SOFA score), severity of disease (APACHE II score) and inotrope therapy (Table 14; for outlier detection and model building see appendix Table 37). A large number of clinical variables were recorded and reduced by

excluding those significantly correlated to each other and by performing outlier analysis for each model (see appendix Figure 31 and Figure 32).

Regression analysis with clinical and demographic data suggested that CAP diagnosis (compared to FP), increasing age and decreasing 25(OH)D levels are significant predictors of an increased probability of death in-hospital and within 6 months of ICU admission. Decreasing mean arterial pressure was the only significant predictor for death during ICU stay.

I next tested the relationship of the variables to elapsed time (days to death) using a proportional hazards model (Cox regression) with the same variables as in the binary logistic regression (Table 14). This confirmed the previous results and indicated that unit increases in age, serum levels of sodium, blood polynucleocytes and diagnosis significantly and independently increase the risk of death. CAP derived sepsis showed a 2.15-fold increased risk of death compared to FP-derived sepsis patients. Conversely, unit decreases in 25(OH)D were an independent and significant predictor of increased risk of death in these patients (4% increase in the risk of death for every 1 nmol/L decrease in 25(OH)D).

**Table 14 Significant predictors of survival outcome after an episode of severe sepsis**

| Survival point | Predictor | B | S.E. | Wald | df | Sig. | Exp(B) | 95% C.I.for EXP(B) | |
|---|---|---|---|---|---|---|---|---|---|
| | | | | | | | | Lower | Upper |
| ICU stay | MAP | -0.19 | 0.06 | 10.93 | 1 | 0.001 | 0.83 | 0.74 | 0.92 |
| | Polynucleocytes | 0.13 | 0.05 | 6.86 | 1 | 0.009 | 1.14 | 1.03 | 1.25 |
| | Age | 0.09 | 0.03 | 7.14 | 1 | 0.008 | 1.10 | 1.02 | 1.17 |
| | Days in ICU | -0.18 | 0.08 | 5.43 | 1 | 0.020 | 0.83 | 0.71 | 0.97 |
| 28 days | Sodium | 0.28 | 0.09 | 9.06 | 1 | 0.003 | 1.32 | 1.10 | 1.58 |
| | MAP | -0.09 | 0.04 | 4.04 | 1 | 0.044 | 0.92 | 0.84 | 0.99 |
| | Polynucleocytes | 0.19 | 0.07 | 6.81 | 1 | 0.009 | 1.20 | 1.05 | 1.38 |
| | Diagnosis (CAP) | 2.90 | 0.87 | 11.08 | 1 | 0.001 | 18.09 | 3.29 | 99.46 |
| | Age | 0.09 | 0.03 | 8.82 | 1 | 0.003 | 1.10 | 1.03 | 1.16 |
| | Sodium | 0.17 | 0.08 | 5.04 | 1 | 0.025 | 1.18 | 1.02 | 1.37 |
| Hospital stay | 25(OH)D | -0.16 | 0.06 | 8.03 | 1 | 0.005 | 0.85 | 0.76 | 0.95 |
| | Polynucleocytes | 0.14 | 0.06 | 5.86 | 1 | 0.015 | 1.15 | 1.03 | 1.28 |
| | Diagnosis (CAP) | 2.53 | 0.78 | 10.53 | 1 | 0.001 | 12.51 | 2.72 | 57.53 |
| Six months | Age | 0.13 | 0.03 | 15.53 | 1 | < 0.000 | 1.14 | 1.07 | 1.21 |
| | Sodium | 0.20 | 0.07 | 7.37 | 1 | 0.007 | 1.22 | 1.06 | 1.41 |
| | 25(OH)D | -0.11 | 0.04 | 8.94 | 1 | 0.003 | 0.90 | 0.83 | 0.96 |
| | Polynucleocytes | 0.12 | 0.05 | 4.80 | 1 | 0.028 | 1.13 | 1.01 | 1.25 |
| | Diagnosis (CAP) | 1.63 | 0.64 | 6.43 | 1 | 0.011 | 5.09 | 1.45 | 17.89 |
| Days to death (Cox) | Age | 0.04 | 0.01 | 9.64 | 1 | 0.002 | 1.04 | 1.02 | 1.07 |
| | Diagnosis (CAP) | 0.77 | 0.34 | 5.22 | 1 | 0.022 | 2.15 | 1.12 | 4.16 |
| | Sodium | 0.06 | 0.03 | 3.94 | 1 | 0.047 | 1.06 | 1.01 | 1.13 |
| | Polynucleocytes | 0.04 | 0.01 | 10.33 | 1 | 0.001 | 1.04 | 1.02 | 1.07 |
| | 25(OH)D | -0.04 | 0.02 | 4.06 | 1 | 0.044 | 0.96 | 0.93 | 0.99 |

Variables that were present in both backward and forward stepwise conditional regression models are shown for predictors with p values <0.050

at the last step (values shown are from forward models). Predicted correct overall percentage tended to increase stepwise and lay at ~87% with Nagelkerke R square tests of ~0.6. B = log odds; SE = standard error; Exp (B)= odds ratio (or hazard ratio for Cox regression); MAP = mean arterial pressure; CAP= community acquired pneumonia (vs. faecal peritonitis). Significance in Cox regression was considered if present on both forward and backward stepwise conditional analyses based on the unadjusted model. Results for the forward model are shown. Days to death were calculated based on the day of ICU admission. Variables included in the final models were: age, bilirubin, days in ICU, sodium, mean arterial pressure, respiratory rate, 25(OH)$D_3$ 1st available sample, polynucleocytes, diagnosis, inotrope therapy, renal replacement therapy and gender.

Vitamin D levels are known to vary with UVB exposure and this is partly dependent on season. I thus ran the same regression analysis to control for seasonal variation. Season alone was not significant but as an interaction term with 25(OH)D was at hospital stay and six month mortality (data not shown). Risk ratios were not materially modified by this term however. As separate terms only 25(OH)D appeared significant. The biological interpretation of this statistical interaction suggests that it is 25(OH)D plasma levels that drives the association to survival outcome following severe sepsis and septic shock. Plasma 25(OH)D in combination with other seasonal factors may additionally influence survival however and complex environmental risk factors may be at play.

***Plasma cathelicidin and 25(OH)D are significantly correlated in patients with severe sepsis***

Antimicrobial peptides are an evolutionarily conserved innate defence mechanism. Cathelicidin-related antimicrobial peptides (LL-37 in humans) are critical against bacterial infection (Vandamme et al., 2012). Several studies have shown a causal relationship between vitamin D and cathelicidin (reviewed in (Hewison, 2011)). I thus measured the levels of this peptide in plasma from the same patients and samples (n=139, median cathelicidin in plasma 35.8 ng/mL, IQR 31.6-39.5 ng/mL, min 25.3 ng/mL, max 99.6 ng/mL; ELISA range 0-100 ng/mL). I correlated these to plasma 25(OH)D and found that they were significantly and positively correlated (rho=0.222, p-value=0.009). In order to assess whether cathelicidin was a confounder for the survival associations I observed earlier I repeated the regression analysis adding cathelicidin. I found that cathelicidin alone nor as an interaction term with 25(OH)D was significant at any survival point. When run as separate terms only 25(OH)D remained significant.

***Plasma 25(OH)D and cathelicidin levels correlate with hundreds of severe sepsis genes***

The transcriptome of leukocytes was measured from the same patient samples to test whether molecular correlates of the association between vitamin D and survival could be detected as well as those of other identified risk factors. Genes expressed in patients with severe sepsis or septic shock were correlated to several biochemical measurements identified in the regression analysis of survival. A Spearman correlation was used for variables important in survival prediction (vitamin D, polynucleocyte count and sodium); thought to be relevant biologically (cathelicidin,

lymphocyte count and monocyte count), or for comparison (bilirubin, urea and creatinine). Microarray gene expression pre-processing and normalisation methods are detailed in chapter two.

I found that hundreds of transcripts were significantly associated to 25(OH)D and also to cathelicidin but not to other plasma molecules (Table 15), including sodium, shown to be a predictor of mortality in this cohort. Many important genes known to play a role in inflammatory processes and in sepsis pathophysiology were significantly associated with vitamin D and also with cathelicidin (p-value<0.05 and false discovery rate <=0.10) including toll like receptors (*TLR1, TLR2, TLR4, TLR5, TLR6, TLR7*, *TLR8, TLR9*), cathelicidin (*CAMP*), interferon members (*IRF2, IRF3, IRF8, IFNGR1, IFNG*), tumour necrosis factor members (*TRADD, FADD, TNFRSF10C, TNFRSF10D, TNFRSF12A, TNFRSF12A, TNFRSF13B, LTBR, TRAF3, TRAF6* amongst others), metalloproteinases (*MMP9, MMP25*), interleukins (*IL1R1, IL7R, IL10RB, IL17RA, IL23A, IRAK1, IRAK3, IRAK4*), NFKB associated molecules (*NFKB1, NFKBIA*), platelet activating factors (PAFAH1B1, PAFAH2), apoptosis related molecules (*BECN1, ATG2B*), and complement (CR1, C1QBP, C5AR1, C1RL, CR2), amongst several others.

**Table 15 Plasma-transcriptome correlations at various FDR cut-offs**

| Plasma molecule | FDR (<=) unique genes | |
|---|---|---|
| | 1% | 5% |
| 25(OH)D | 868 | 1848 |
| Cathelicidin | 1107 | 1980 |
| Bilirubin | 0 | 21 |
| Urea | 0 | 0 |
| Sodium | 0 | 0 |
| Creatinine | 0 | 0 |

A Spearman rank correlation was used to analyse the association of various measurements in GAinS patients with gene expression from the leukocytes of the same individuals. Significance was considered with a T-test p-value <0.05 at various false discovery rates as presented in the table. The leukocyte whole genome transcriptome was analysed using Illumina arrays as described in chapter two. Microarray data was normalised using the distance weighted discrimination method and analysed using BeadStudio software (V3.1, Illumina), lumi on R (Du et al., 2008) and Limma on R (v. 2.10.1, Bioconductor v. 2.5). Outliers and probes with low detection confidence were removed before further analysis.

***Severe sepsis genes are regulated by vitamin D***

In order to test whether plasma 25(OH)D correlated genes may actually be regulated by vitamin D I measured enrichment of VDR binding sites near these genes. I used all genes expressed in sepsis patients as background and found that VDR binding sites are present near plasma 25(OH)D correlated genes more than expected by chance

(~1.5 fold enrichment, corrected p-value (q-value) <0.05). Moreover, I found that this enrichment is specific to VDR as compared to glucocorticoid receptor and that it is specific to VDR binding in leukocytes (B cells and monocytes) as compared to VDR binding in colonic cells (Table 16).

**Table 16 Leukocyte genes expressed in patients with severe sepsis are likely to have VDR binding sites nearby**

| Track | Annotation | Observed | Expected | SD | Fold | P-value | Q-value |
|---|---|---|---|---|---|---|---|
| Glucocorticoid receptor binding | Sepsis-25OHD | 33462 | 35013 | 4740 | 0.96 | 0.37835 | 0.40537 |
| | Sepsis-cathelicidin | 52188 | 54066 | 5796 | 0.97 | 0.37743 | 0.40537 |
| VDR binding colonic cells | Sepsis-25OHD | 24378 | 24930 | 3550 | 0.98 | 0.44479 | 0.45765 |
| | Sepsis-cathelicidin | 38527 | 40027 | 4447 | 0.96 | 0.3745 | 0.40537 |
| VDR binding lymphoblastoid cells | Sepsis-25OHD | 155297 | 87650 | 9330 | 1.77 | 0.00001 | 0.00004 |
| | Sepsis-cathelicidin | 164401 | 134869 | 11368 | 1.22 | 0.00575 | 0.01437 |
| VDR binding monocytes | Sepsis-25OHD | 228488 | 143631 | 14713 | 1.59 | 0.00001 | 0.00004 |
| | Sepsis-cathelicidin | 281762 | 209981 | 17328 | 1.34 | 0.00001 | 0.00004 |

Analysis of overlap between genome-wide binding sites and genomic intervals near the transcriptional start site of genes expressed in patients with severe sepsis or septic shock was carried out with GAT (see Chapter 2:

Materials and Methods) (Heger et al., 2013). Binding sites were derived for glucocorticoid receptor (Reddy et al., 2009), VDR in colonic cells (Meyer et al., 2012), VDR in lymphoblastoid cells (Ramagopalan et al., 2010) and VDR in THP1 cells (monocyte like cell line) (Heikkinen et al., 2011). Genes with expression values that correlated with cathelicidin or vitamin D with a p-value<0.05 and FDR<0.01 (Spearman) were used. All genes expressed in the sepsis patient cohort were used as background. Results indicate transcription factor binding enrichment near vitamin D (or cathelicidin) correlated genes over all expressed genes in the same patients.

***Vitamin D mediated antimicrobial mechanisms are present in severe sepsis patients***

I carried out pathway analysis on genes correlated to plasma 25(OH)D or cathelicidin and found signatures of known vitamin D mediated antimicrobial responses in both sets including interferon gamma (IFNG), Toll like receptors (TLR) and nuclear factor kappa B (NFKB) mechanisms. Surprisingly, I find that the most significant pathways associated to 25(OH)D are high level regulatory functions such as protein synthesis and cell survival while cathelicidin correlated genes involve classic inflammatory processes. Interestingly, genes correlated to cell counts (polynucleocytes, lymphocytes and monoctyes) did not result in any significant pathways (see appendix Table 39, Table 40 and Table 41 for full list of enriched pathways for genes significantly correlated to vitamin D, cathelicidin or polymorphonuclear cells; lymphocytes and monocytes correlated pathways or genes not shown).

***Gene expression in severe sepsis survival does not reveal differences***

In order to better understand what the global transcriptomic differences were between survivors and non-survivors, the differentially expressed genes between groups were compared. Surprisingly, no significant differences were found for genes with >1.5 fold-change and corrected p-values <0.05. Although this is counter-intuitive, it is in agreement with a previous study comparing complicated and uncomplicated cases in burn and severe trauma patients (Xiao et al., 2011). Given these results, I was next interested in testing whether combining clinical and molecular data could reveal patient differences that were otherwise not evident.

***Molecular phenotyping identifies individuals at increased risk of early death***

As plasma vitamin D was a significant predictor of survival and genes correlated to it appeared to be biologically relevant (Table 15, Table 16 and appendix), I integrated this with other survival predictor data to identify sub-groups of patients. Although pathway analysis was not significant for any of the cell counts, I used polynucleocyte count as it was a significant predictor of mortality and showed a high number of genes correlated to it. Blood sodium levels, lymphocyte count and monocyte count were either not survival predictors, did not have correlated genes or did not show evidence of biological relevance using pathway analysis.

A graphical representation of expressed genes correlated to plasma 25(OH)D (adjusted p-value <0.05) was first generated by performing hierarchical clustering and arranging results according to survival and plasma vitamin D correlation (Figure 27 and appendix Figure 34). A clear pattern of clusters became visually evident with four groups of patients.

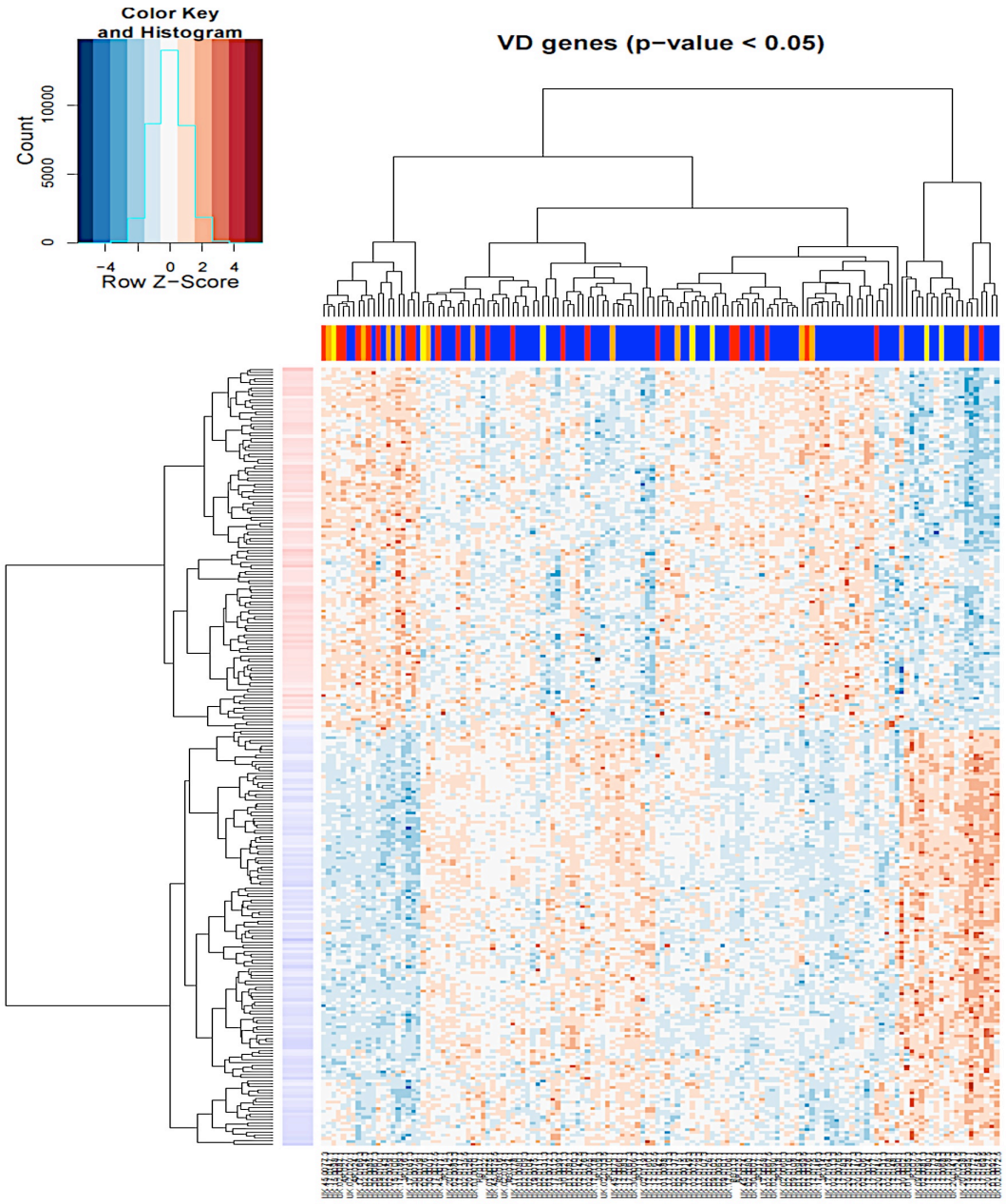



**Figure 27 Vitamin D correlated genes reveal patient sub-groups.**

The heatmap represents hierarchical clustering of genes expressed in patients with severe sepsis and septic shock. Genes that correlated to both plasma 25(OH)D and polynucleocyte count were used. Results on the vertical axis were arranged according to correlation with polynucleocyte count (red,

positive; blue, negative, Spearman rho). Gene expression values represent Z-scores standardised across samples and relative to mean expression (red = expression above the mean; blue expression below the mean). On the horizontal axis patients were arranged according to survival (red = death during ICU stay; orange = discharged alive from ICU but died during hospital stay; yellow = discharged alive from hospital but died within six months; blue = alive at end of study [six months]).

To quantify this, the enrichment of non-survivors in cluster one (leftmost on the heatmap, Figure 27) vs. other clusters were tested. I found a significantly higher number of non-survivors present in patient cluster one compared to others and that an important proportion of these patients died within 10 days of ICU admission ('individuals at high risk of early death', Figure 28)(Fisher's exact test, p-value <1.00E-17).

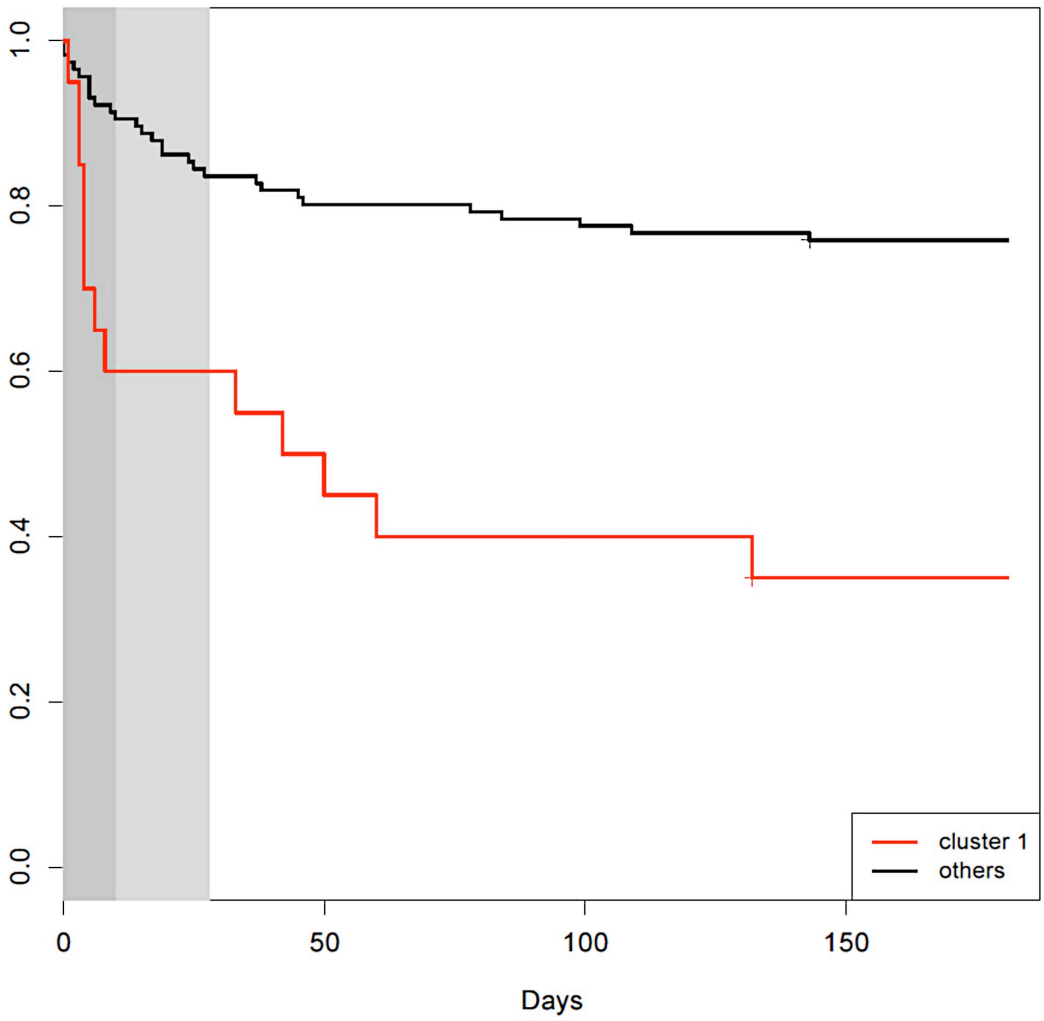



**Figure 28 Molecular phenotyping identifies patients at high risk of early death.**

The Kaplan-Meier plot compares cluster one (left-most in Figure 27) vs. the three remaining clusters. Y-axis = fraction of patients alive; x-axis = days to death; dark gray = 10 days, light gray = 28 days; total number of patients belonging to cluster 1 = 20 [total non-survivors = 13]; to other clusters n = 116 [total non-survivors = 29].

The high-risk individuals were also found to have genome-wide differences in gene expression when compared to others. The expression of over 2500 genes was significantly different between the clusters (cluster one vs. others; >1.5 fold change, adjusted p-value <0.05) (see appendix). Finally, to assess consistency we measured the relationship to vitamin D of genes differentially expressed between patient clusters by testing the enrichment of plasma 25(OH)D correlated genes previously identified. The presence of over 300 vitamin D genes was found in the patient cluster differentially expressed gene set and these vitamin D genes were significantly more enriched in genes expressed in high-risk individuals (cluster 1) vs. other patients (adjusted p-value <0.05, Fisher's exact test).

***Molecular identification of high risk patients is not clinically evident***

I investigated whether clinical variables already collected could distinguish patients at high risk. I tested all variables available as for the initial univariate analysis but found that neither APACHE, SOFA or diagnosis were different between the clusters while, as expected, plasma vitamin D, polynucleocyte count and cathelicidin were significantly different (Mann Whitney U or Fisher's exact test, two tailed exact p-values <0.05). Interestingly, gender and the Charlson index showed significant differences with more females and individuals with lower comorbidity at high risk for early death (Mann Whitney U or Fisher's exact test, two tailed exact p-values <0.05).

## Discussion

I have carried out an observational study interrogating the levels of 25(OH)D in plasma in patients with severe sepsis or septic shock and found that this group of patients is severely deficient. I correlated these values to available clinical data and modelled survival outcome showing that decreasing levels of 25(OH)D significantly increase the likelihood of death. I next showed that these observations were reflected at the molecular level by correlating biochemical and clinical parameters with gene expression. This integrated approach, based on plasma vitamin D and the leukocyte transcriptome, revealed a subset of patients at increased risk of early death that was not evident from clinical parameters. The group of high-risk individuals was significantly enriched for female patients with a lower index of comorbidity who died within 10 days of ICU admission.

Several studies have looked at vitamin D deficiency in hospital and critical care patients. A recent investigation did not find an association of 25(OH)D at 28 day mortality following sepsis, with only age appearing significant, in agreement with the results in this chapter (Cecchi et al., 2011). A smaller sample size was used and long-term mortality was not measured. Another recent study found that, in a cohort of 2399 ICU patients, vitamin D deficiency and insufficiency prior to hospital admission are significant predictors of short and long-term mortality and increased risk of blood culture positivity (Braun et al., 2011). To my knowledge, this study is the first to address vitamin D levels prior to hospitalisation. Data obtention and analysis were carried out retrospectively through a computerised registry however, and why this

particular group of patients had 25(OH)D measured is not clear as it is not routine clinical care. It is possible that more investigations were ordered in patients who had more severe disease or other underlying pathology. Despite limitations, this study and others highlight the growing evidence of involvement of vitamin D in mortality in critically ill patients.

Adding to these previous studies I show that vitamin D deficiency in sepsis is highly prevalent and severe, that plasma 25(OH)D predicts survival outcome, that this is clearly reflected at the molecular level and that comprehensive profiling can identify sub-groups of patients with different survival outcomes. The multivariate regression model for survival predicts that every 10 nmol/L decrease in 25(OH)D plasma levels increases the odds of death by 0.01 to 0.07. The change is relatively small but significant across different survival time-points, including in-hospital, at 6 months and in a Cox regression using days to death as a continuous variable. In a treatment scenario this would equate to a 1 to 17% decrease in mortality risk when comparing a severely deficient patient to one with 'optimal' vitamin D levels (<25 vs. >75 nmol/L).

This study is keeping with previous clinical findings of vitamin D deficiency in ICU patients (Braun et al., 2011; Braun et al., 2011; Cecchi et al., 2011; Ginde et al., 2011; Jeng et al., 2009) and there is evidence from animal and human studies that administration of vitamin D is safe and may be beneficial in acute states (Amrein et al., 2011; Horiuchi et al., 1991; Mata-Granados et al., 2010; McKinney et al., 2011).

Integrating clinical and molecular variables allowed identification of patients at increased risk of early death and showed that vitamin D associated genes significantly contributed to differentially expressed genes in high-risk individuals. I observed clear signs of known vitamin D mediated antimicrobial mechanisms active in our cohort, including involvement of interferon gamma and cathelicidin. Genes and pathways that may serve as candidates for further studies aiming to dissect disease mechanisms in inflammatory conditions are highlighted. *In-vivo* evidence that vitamin D and cathelicidin interact and may be an important mechanism in sepsis is also provided.

This study has important limitations however and may be confounded due to the complex nature of both sepsis and the metabolism of vitamin D. A potential confounder for measurements of 25(OH)D in plasma is haemodilution. A recent study with a small sample size (n=19) looked at shifts in circulating vitamin D in critically ill patients treated with intravenous fluid therapy. They found low values of vitamin D due to haemodilution that took 24 hours to resolve (Krishnan et al., 2010). However, a recent publication comparing patients with sepsis and trauma found a significant difference in vitamin D levels on admission to ICU (samples drawn <24 hours) but no difference in markers of fluid therapy (Cecchi et al., 2011). Nutritional therapy that includes cholecalciferol may contribute to these differences and fluid resuscitation, diet or plasma proteins (e.g. albumin to normalise or compare trends) were not available for our cohort.

The GAinS cohort is representative of CAP and FP derived sepsis patients in the ICU in the UK and could thus be generalised to these types of patients. Few studies have

addressed transcriptome-metabolite/protein correlations in humans. The present analysis revealed hundreds of genes associated to vitamin D and pathway analysis showed agreement with previous knowledge of potential mechanisms in sepsis (Barnay-Verdier et al., 2011; Zanoni et al., 2009). Genes associated to vitamin D and to other predictors of sepsis outcome were highlighted by this approach and may serve as candidates for further studies.

Although I am not able to probe causal relationships with the current data, they raise the question as to whether vitamin D is a risk factor in sepsis incidence and mortality. This question should be addressed in a prospective cohort. Importantly, restoring levels in critically ill patients deficient in vitamin D may be beneficial regardless of causal factors. This study suggests that small changes in the levels of vitamin D may have important effects. The safety and efficacy of vitamin D administration has been tested in pilot trials with small sample sizes in groups of ICU patients with various diagnoses (Amrein et al., 2011) (McKinney et al., 2011). However, long-term safety and benefits must be established and further replication is necessary to address possible confounders.

Given the findings made here and those of other groups these results suggest that prior low levels or acute consumption of vitamin D significantly contribute to a dysregulated systemic inflammatory state that eventually leads to death. Work is required to ascertain the association of vitamin D deficiency to sepsis and survival through a cohort specifically designed with these outcomes. This may then enable research in disease mechanisms and trials of clinical applications.

***Clinical and molecular correlates of vitamin D and sepsis survival:***

- Patients with severe sepsis or septic shock have a high prevalence of severe vitamin D deficiency;
- Vitamin D levels are significant and independent predictors of long term mortality in patients suffering from severe sepsis or septic shock;
- Vitamin D levels correlate to many genes expressed in leukocytes in septic patients;
- Vitamin D associated genes belong to pathways essential in immune and inflammatory responses that are significantly enriched in septic patients.
- Molecular phenotyping using vitamin D and other outcome risk factors identify patients at increased risk.

**Figure 29 Specific conclusions from chapter five.**

# CHAPTER 6: GENERAL DISCUSSION, CONCLUSIONS AND FUTURE WORK

## Importance of vitamin D in immunity, inflammation and disease

Over the last decade methods to reliably identify genetic risk in disease have been developed, particularly using genome-wide association studies (Burton et al., 2007). Our understanding of the genetic basis in complex conditions has thus grown in recent years and specific disease-associated loci in particular populations have been clearly defined. Genetic risk is not the sole contributor to disease risk however and evidence indicates that both genetic and environmental factors are causal in MS and other diseases (Giovannoni et al., 2007; Gwinn et al., 2009).

The importance of vitamin D in bone physiology and disease has been known for decades but its wider role in immunity and inflammation is coming to light (Holick, 2007). Vitamin D plays a key role in multiple physiological functions and has now been extensively associated with disease. Notable disorders include the well-studied bone alterations but also autoimmune, cancer and infectious diseases. VDR, the molecular mediator of vitamin D actions, is crucial in physiology and disease, as genetic associations have clearly linked it to multiple traits. However, the exact molecular mechanisms as to how VDR exerts its genetic control over a vast number of genes and cell types is not fully understood and is currently under intense scrutiny.

Novel applications based on the advancement of sequencing technology allow characterisation of transcription factor profiling in diverse species and populations

and provide a powerful unbiased approach to further our knowledge of the molecular basis of complex disease. Methods interrogating gene environment interactions have revealed the vast number of genes regulated by vitamin D and recent studies are providing the molecular basis for the association of vitamin D to MS and other diseases (Ramagopalan et al., 2009; Wang et al., 2010).

The association between vitamin D and MS is supported by epidemiological and genetic observations (Giovannoni et al., 2007; Ramagopalan et al., 2011). It is possible that early life events, genetic risk and inheritance of acquired states increase the risk of dysregulated immune responses that become apparent under environmental stress in later periods (Ascherio et al., 2012). The evidence for vitamin D as a causal factor in MS is strong although many questions remain unanswered. The mechanism(s) as to how vitamin D exerts its genetic control and its molecular association with disease are not fully understood however.

In this thesis I addressed some of these questions using epidemiological and functional genomics methods. I showed that disease-associated variants disrupt VDR binding, that vitamin D contributes significantly to immune and inflammatory responses and that vitamin D deficiency is an important risk factor in inflammatory conditions such as sepsis. The overarching thesis is that vitamin D is an important modulator of immunity and that its regulation, interaction with genetic variation, blood levels and actions are important in disease. I used two relevant conditions to exemplify this, MS and sepsis. The association between vitamin D and MS is better understood and evidence suggesting vitamin D as a causal factor has been shown

recently (Munger et al., 2006; Ramagopalan et al., 2011; Sawcer et al., 2011). The association between sepsis and vitamin D is tenuous although extrapolation from other intensive care conditions and sub-group analysis offers some insight (Braun et al., 2011; Braun et al., 2011).

## Genetic variation, VDR binding and autoimmune disease

MS serves as a paradigm in autoimmune research in gene-environment interactions and is a condition where the evidence for vitamin D as a causal factor is strongest. The molecular mechanisms behind this association are becoming increasingly understood. Evidence of vitamin D regulation of MS associated genes provides further rationale for use of vitamin D as an agent for prevention and/or treatment. Haplotype specific responses induced by vitamin D independently or through complex interactions could aid our understanding of how important risk variants contribute to disease mechanisms and respond to therapy in MS, sepsis and other diseases.

Genome-wide analyses showed that binding of VDR occurs near MS and other disease-associated variants more than expected by chance, indicating that vitamin D regulation of gene expression may be widespread and important in disease. I provide new evidence of disease relevant genetic variation at VDR binding sites. I show that a single nucleotide polymorphism within a VDR interval near the MS gene *IL2RA* is associated to MS susceptibility in a Canadian cohort. Additionally, I show two of the first examples where autoimmune-associated variants found within VDR motifs disrupt binding.

These results provide evidence that variation at VDR binding intervals may be relevant in autoimmune pathophysiology. The experiments performed here have important limitations but taken collectively indicate that vitamin D may be an

important modulator of disease relevant genes and that genetic variation impacts its effects.

### VITAMIN D AND MULTIPLE SCLEROSIS GENETIC SUSCEPTIBILITY: IMPLICATIONS FOR MS CAUSALITY AND PREVENTION

Many genes implicated in MS susceptibility, including the main locus HLA-DRB1*1501, are regulated by vitamin D (Ramagopalan et al., 2010; Ramagopalan et al., 2009). It is likely that vitamin D plays more than one role in more than one point in time and space. A picture is emerging pointing towards at least three spatio-temporal effects of vitamin D. Deficiency during pregnancy either during the first or second trimester may affect thymic development and levels of T cell auto-reactivity (Taylor et al., 2010).

Secondly, chronic exposure to low levels during childhood and up to adolescence may cause immune system dysregulation by altering peripheral tolerance and priming immune cells for specific responses, possibly through chromatin landscape modifications that alter gene expression patterns and later contribute to disease.

Thirdly, vitamin D may act during disease by protecting against relapses through a shift from a Th1 to a Th2 type response (Lemire, 1995; Mattner et al., 2000), inhibiting a Th17 response (van Etten et al., 2007; Xue et al., 2003), by increasing the function of regulatory B and T cells (Adorini et al., 2009; Gregori et al., 2001; Heine et al., 2008; Penna et al., 2005; Smolders et al., 2009) and suppressing plasma cell differentiation and antibody production (Lemire et al., 1984; Linker-Israeli et al., 2001); amongst other effects (Baeke et al., 2010). T helper cell differentiation, dendritic cell maturation and cross talk between dendritic cells and natural killer cells

may be relevant pathways to consider in MS pathophysiology given previous evidence and implicated genes. These observations are supported by immunological abnormalities seen in MS patients (Gandhi et al., 2010).

Given an estimated one billion-plus people worldwide with inadequate vitamin D levels (Holick, 2007) and multiple observations associating vitamin D deficiency to disease, a question now is how to examine the efficacy and effectiveness of intervention?

Studies of disease susceptibility regarding month of birth and parental transmission of risk haplotypes implicate early-life periods and possibly ancestral effects. It is likely during this time period when vitamin D and MS gene interactions are most important and set the stage for autoimmunity. Epigenetic inheritance appears to exist and would magnify the public health implications if relevant in MS (Grossniklaus et al., 2013).

Migration data and variability in relapse rates support effects in adolescence and even into adulthood and during disease (Ascherio et al., 2012). These observations are likely caused by immune system dysregulation through the influence on peripheral tolerance and effector mechanisms (Simon et al., 2012). Future work could address the functionality of associated genetic variants and vitamin D mediated mechanisms. Importantly, studies of vitamin D supplementation are warranted to address efficacy and efficiency in MS patients and relatives at high risk.

## Integrative analysis of genetic variation, gene expression and plasma vitamin D

The explosion in data production in the genomics era has brought a rapid expansion of techniques and possibilities that allow unparalleled study of organisms at various levels. Genome-wide analyses of genetic variation, gene expression and other layers of information that can be correlated to phenotypes permit detailed characterisation. Oncological studies have led the way in integrative genomics methods and have provided important insight into disease mechanisms and patient phenotyping (Kandoth et al., 2013). Few integrated studies outside of this area have been performed however and none related to vitamin D or sepsis to my knowledge. Data analysis and interpretation bottlenecks are now important limitations however and methods of integrating diverse sources of data are currently being developed. Hypothesis driven research with insightful interpretation remain essential.

In chapters four and five I use an integrative genomics approach to characterise the interplay between genotype, transcription, plasma vitamin D and phenotype. I show that genetic variation regulates blood levels of vitamin D and that vitamin D in turn is associated to cell specific expression. I analysed the distribution of 25(OH)D in plasma of healthy individuals and found that, in agreement with published data, a considerable proportion of Caucasian British adults are vitamin D deficient or insufficient. SNPs associated to low levels of vitamin D were also confirmed in this volunteer cohort. Finally, expressed genes from monocyte and B cell transcriptomes that correlated to plasma levels of 25(OH)D showed stark cell specific differences.

Sepsis is a devastating disease with unmet needs. It is derived from a systemic inflammatory dysregulated response to infection. Nearly 30 clinical trials in the last two decades have been negative and accurate diagnosis and prognosis remains challenging. Few studies have addressed the role of vitamin D in sepsis. I postulated that vitamin D would play an important part in this disease and that integrating clinical, biochemical and molecular data could provide important insight. I first showed in a cohort of patients with severe sepsis or septic shock that vitamin D deficiency was highly prevalent and severe. This is in accordance with previous studies in hospitalised and critically ill patients although the severity and prevalence were markedly higher in the GAinS cohort (Braun et al., 2011; Braun et al., 2011; Moromizato et al., 2013).

I next showed that vitamin D levels are significant and independent predictors of long-term mortality in sepsis patients after correcting for important clinical and biochemical variables. These findings showed that low levels of vitamin D appear to be significant determinants of sepsis outcome. The next question was whether these clinical observations were reflected at the molecular level. I performed a similar analysis as in the healthy individuals and correlated vitamin D levels to leukocyte gene expression in sepsis patients. This showed that many genes were highly correlated and that they belonged to pathways essential in immune and inflammatory responses significantly enriched in severe sepsis patients. Several of these pathways have been previously implicated in sepsis physiopathology (Bosmann et al., 2013).

Importantly, this molecular phenotyping and integrated analysis using vitamin D, other outcome risk factors and gene expression identified patients at increased risk of early death. This sub-group of patients was not identifiable by clinical characteristics such as the prognostic score (APACHE) or organ failure assessment (SOFA). The high-risk group was characterised by lower levels of vitamin D and cathelicidin, higher numbers of circulating PMNs, a particular leukocyte transcriptome signature, female gender, a lower comorbidity score (Charlson index) and death within 10 days of ICU admission.

The results here raise the question of whether vitamin D treatment may be beneficial in this group of patients and also support the design of prospective cohorts to interrogate causality. Additionally, they suggest that deeper phenotyping, including molecular and genetic data, can identify patient heterogeneity that is not otherwise evident and which has prognostic implications.

Cancer research has pioneered integrative research using genomics approaches in the last decade and progress has been made in identifying druggable targets and in more accurately classifying patients (Hudson et al., 2010; Kandoth et al., 2013). Research using this approach has lagged in other areas but the clinical utility is tangible. Research in vitamin D has seen an upward trend in recent years but is largely observational. Investigations with prospective cohorts and using randomised trials are necessary to address causality and the value of treatment with vitamin D in the many conditions to which it has been associated. Although the investigations in healthy individuals and in patients with severe sepsis and septic shock are observational, they

shed light on vitamin D biology and on disease mechanisms. These studies have highlighted important molecules, pathways and approaches that may serve in translational research.

# Clinical implications of vitamin D deficiency in severe sepsis

## ANALYSIS OF CAUSALITY USING HILL'S CONDITIONS

Although there are no set criteria to determine causality, conditions have been suggested and the most used in epidemiology are those listed by Austin Bradford Hill (Hill, 1965). These are a set of aspects that can guide the assessment of causality once an association is clearly established and is statistically unlikely to be due to chance or to be biased. Sufficient evidence, although observational and largely from mixed cases of intensive care or hospitalised patients, has now accumulated and allows an initial analysis of causality of vitamin D deficiency in critical conditions.

### *Strength of the association and consistency between studies*

The odds ratio (OR) I identified in the regression analysis indicated that the strength of the association was moderate. Hazard ratio analysis showed a 4% increase in risk per unit decrease of plasma 25(OH)D (1 nmol/L). The effect was independent and statistically significant for lowering plasma vitamin D (25(OH)D) and increasing risk of death following severe sepsis or septic shock. A decrease of 50 nmol/L (from 75 to 25 nmol/L, sufficiency to deficiency) of 25-hydroxyvitamin D would increase the risk of death by 2-fold. The association was consistent for mortality in-hospital and at 6 months after ICU discharge as well as when using days to death as a continuous variable. This study had comparable results to previous observations of vitamin D deficiency in ICU and septic patients.

Three studies have measured vitamin D levels in sepsis patients. Although there may be a publication bias, only one found negative results (Cecchi et al., 2011). It assessed

one survival time-point, 28-day mortality, and was likely underpowered as it only measured levels in 92 septic patients. An earlier investigation in a small number of intensive care patients with and without sepsis (n=24 and 25, respectively) found that plasma 25-hydroxyvitamin D levels were significantly lower in both groups when compared to healthy controls (Jeng et al., 2009). It did not assess mortality outcomes however. A third study in patients with suspected infection in the emergency department found that those with lower levels of plasma vitamin D (<75 nmol/L vs. > 75 nmol/L) had a significantly higher risk of developing septic shock and higher SOFA and APACHE scores 24 hours later (Ginde et al., 2011).

Other studies have found consistent ORs ranging between ~1.2 and ~1.9 when comparing groups with higher versus lower vitamin D levels in large cohorts of mixed diagnostic cases of hospitalised or critical care patients (Braun et al., 2011; Braun et al., 2011; Lange et al., 2013; Moromizato et al., 2013). One of these studies measured vitamin D seven days before or after the start of critical care in 1,325 patients and found an OR of 1.94 (adjusted, 95% CI 1.18-3.20; p = 0.01) for prediction of mortality at 30 days after intensive care when comparing patients with 25-hydroxyvitamin D ≤ 37 nmol/L versus those with ≥ 75 nnmol/L (Braun et al., 2011). Ratios for 90 days, 365 days and in-hospital mortality were similar.

A recent study by the same researchers found that vitamin D deficiency prior to hospital admission significantly predicted short and long-term mortality (Braun et al., 2011). Adjusted odds ratios using the same thresholds as above were of 1.69 (95% CI 1.26-2.26, p < .0001; total n=2,399) and similar for 90 and 365-day mortality. A

recent investigation with a similar design analysed 23,603 hospitalised patients and found comparable results for the same thresholds and survival points (Lange et al., 2013). A fourth study also measured 25-hydroxyvitamin D deficiency prior to hospital admission and found it to be a significant predictor of critically ill sepsis and of increased mortality following intensive care initiation in those with 25(OH)D <75 nmol/L (Moromizato et al., 2013).

These studies are encouraging but have serious limitations. All studies were performed retrospectively, from the same geographic area and hospitals (Boston, USA) and duplication bias was not clearly addressed. Although they are consistent in direction and magnitude with the results presented in chapter five, further independent validation in other populations and in specific conditions is necessary.

***Biological gradient and temporality***

The range of plasma vitamin D values in the GAinS cohort was short but suggested that small changes had significant effects on survival. Additionally, I used plasma vitamin D and an outcome measure (days to death) as continuous variables indicating a biological gradient. Other studies in critical care and hospitalised patients, described above, have compared survival outcomes using vitamin D thresholds (Braun et al., 2011; Braun et al., 2011; Lange et al., 2013; Moromizato et al., 2013). Gradients in these studies are suggested by odds ratios that were larger for vitamin D deficient patients than for those with insufficiency when compared to vitamin D sufficient patients.

These studies also had stronger ORs for short-term mortality. Although my study did not show an association for in-ICU or 28-day mortality, this may be due to lack of power. I was not able to assess temporality in the GAinS cohort but studies using large cohorts determined that pre-hospital admission levels of 25-hydroxyvitamin D were significantly correlated to outcome (Braun et al., 2011; Lange et al., 2013). These are the first to analyse pre-admission levels of vitamin D in intensive care and hospitalised patients.

A study in patients with suspected infection, mentioned earlier, found vitamin D deficient patients to be more likely to develop sepsis and more severe disease 24 hours later (Ginde et al., 2011). Although consumption in acute states has been described (Bang et al., 2011), plasma 25-hydroxyvitamin D has a long half-life (Jones, 2008; Mawer EB, 1971). Pre-disease investigations with satisfactory ascertainment and adequate control of confounders are yet to be carried out and would be necessary before any conclusions can be made.

***Specificity of the association and reasoning by analogy***

The association of vitamin D and outcome from sepsis does not seem specific. Susceptibility, outcome and response to treatment in many diseases have now been associated with varying levels of vitamin D (Holick, 2007). Vitamin D deficiency causes rickets and osteomalacia and appears to also significantly contribute to MS susceptibility (Beecham et al., 2013; Ebers, 2008; Ramagopalan et al., 2011). Inflammatory and autoimmune conditions have far less evidence although mechanisms of action in immune regulation by vitamin D may be a common factor.

This may be true for sepsis where a dysregulated inflammatory state lies behind the often-fatal physiopathological response.

Interestingly, as with MS, seasonal variation significantly influences sepsis incidence and mortality with higher rates in winter compared to the fall in the northern hemisphere (Danai et al., 2007). A latitude gradient was also described in the same study with higher rates of incidence and mortality in the northeast of the USA (Danai et al., 2007). Replication in distinct geographic regions and in the southern hemisphere is required however and socioeconomic and ethnic backgrounds need to be fully accounted for.

***Biological plausibility and coherence with sepsis mechanisms***

The discovery that VDR and vitamin D enzymes are expressed ubiquitously and that vitamin D is an important regulator of immune and inflammatory responses have raised the possibility that vitamin D deficiency could influence disease susceptibility and outcome. Clinical and epidemiological observations of vitamin D deficiency have led to molecular associations in some diseases. In this study I presented observational evidence of both for severe sepsis and septic shock. Pathway analysis revealed significant enrichment of genes associated to plasma vitamin D and known to have important roles in sepsis physiopathology. The vitamin D mediated interferon gamma-antimicrobial peptide response pathway and other known antibacterial functions regulated by vitamin D make it a coherent and likely agent for both sepsis susceptibility and outcome modulation (Fabri et al., 2011; Hewison, 2011).

***Experimental evidence: vitamin D as possible adjunct therapy in intensive care patients***

There are a number of potential mechanisms through which vitamin D could act as it is known to promote an anti-inflammatory state while enhancing antimicrobial actions (Fabri et al., 2011; Hewison, 2011; Jeng et al., 2009; Moller et al., 2007). Molecular analyses in clinical observational studies can provide functional links and improve biomarker search and give insight into possible mechanisms of disease. If observational studies have not been confounded, it is possible that restoring levels in critically ill patients deficient in vitamin D may be beneficial regardless of initial diagnosis. Administration of vitamin D, most likely as adjunct therapy and as $25(OH)D_3$, appears feasible in deficient patients given its safety profile and proved importance in calcium homeostasis.

Recent trials with small sample sizes in intensive care patients have shown that administration of vitamin $D_3$ (cholecalciferol, pre-$25(OH)D_3$; the inactive, un-hydroxylated vitamin $D_3$ form) is safe and may be beneficial. A pilot phase I trial in 25 ICU patients showed the safety of a single high oral dose of cholecalciferol in the short term (Amrein et al., 2011) and there is evidence from animal and human studies that administration of vitamin D is beneficial (Horiuchi et al., 1991; McKinney et al., 2011). Two small trials in ICU patients gave initial evidence that vitamin D normalises levels, reduces ICU stay and associates with survival (Mata-Granados et al., 2010; McKinney et al., 2011).

The stage of illness and the environmental milieu, particularly microbial interactions, may influence the effect of vitamin D. Trial designs for vitamin D may have many

pitfalls as the type of vitamin D, dose, available formulations, oral administration in patients with intestinal absorption compromise, metabolism passes at liver and kidney in failing organs, availability of plasma 25-hydroxyvitamin D tests and several other issues may confound results. Randomised trials with carefully phenotyped individuals are required to adequately assess the value of vitamin D supplementation in sepsis and critically ill individuals.

### MOLECULAR PHENOTYPING IN SEVERE SEPSIS PATIENTS

Although much is understood, sepsis mechanisms remain poorly detailed. This has been reflected by the lack of progress in developing targeted therapeutics (Cohen et al., 2012). Clinical trials in severe sepsis and septic shock have been disappointing and treatment is largely based on supportive care (Dellinger et al., 2013). Integrative genomics may offer an approach to better understand basic mechanisms of disease, identify druggable targets and allow improved patient classification which is not evident using clinical parameters. Cancer research has lead the way in using genetic and genomic information to drive molecular classification and personalised therapeutics with early cases of success (Chabner, 2011; Hudson et al., 2010). In chapter five I showed that integrating risk factor identification with biochemical, transcriptomic and clinical parameters in sepsis patients yielded insight that is potentially translatable.

# Conclusions and future work

The combination of molecular and epidemiological approaches can provide information that may be clinically relevant. This approach may lead to clinical translation in terms of improved diagnostic classification and treatment. The results in this thesis provide links between molecular and clinical observations that are relevant to our understanding of vitamin D biology in health and disease. Many areas remain open and important questions continue to arise: Is VDR a global genome organiser? Does it regulate environmental inputs on a global physiological scale? What is the functional interplay between genetic diversity and vitamin D? What epigenetic mechanisms are important in VDR mediated transcriptional regulation? Why does a variable environmental factor play such an important role in immunity? What are the exact mechanisms of vitamin D regulation of cell specific responses? Importantly, can vitamin D help prevent or treat autoimmune and inflammatory diseases?

Further studies are necessary to understand the time-points, cell types, molecular targets and mechanisms involved in the actions of vitamin D. Epigenetic factors, inheritance and changing epidemiological patterns are poorly understood. The interaction between vitamin D, EBV and genetic variants may also help explain disease mechanisms and allow patient stratification for risk and therapy. Further evidence of vitamin D deficiency as a causal factor, its molecular targets in MS and its prospect as a therapeutic and preventative agent are questions that warrant further study and are under current investigation. Several clinical trials have been completed and more are under way to understand the full extent and value of vitamin D

supplementation on disease progression in MS and other diseases. Large clinical trials and prospective studies are necessary to assess the true benefit of vitamin D, ideally as a preventive agent. Careful study design is increasingly pertinent. Timing of exposure, dosage and trans-generational effects are some of the many important questions that need to be addressed in vitamin D and MS research.

The evidence for the association of low vitamin D plasma levels in other inflammatory conditions is far less strong. Only observational studies and retrospective studies have been carried out to understand the role of vitamin D in sepsis and causal links remain highly uncertain. Many confounders need addressing and molecular studies are necessary to understand possible mechanisms. Genetic associations have not been observed and the value of vitamin D supplementation in sepsis is only beginning to be explored. Further studies to ascertain the importance of vitamin D, disease mechanisms and potential benefit as a preventive and/or therapeutic agent seem justified given the high prevalence of its deficiency and number of disease associations.

**General conclusions**

- Regulation of vitamin D is complex and likely tissue specific;
- Genetic variation has an important impact on regulation of and by vitamin D;
- Regulation of biological processes by vitamin D seems widespread, tissue specific and important in the immune system;
- Vitamin D levels in plasma are low in the general population as well as in sepsis patients.
- Vitamin D is associated to mortality and is an important biomarker in sepsis patients;
- Integrated analysis using diverse layers of biological information can provide insight into vitamin D biology, disease mechanisms and may be useful in patient classification.

**Figure 30 General conclusions derived from this thesis.**

# APPENDIX

## Chapter 1: Literature review

### MULTIPLE SCLEROSIS: A COMPLEX DISEASE INFLUENCED BY VITAMIN D

**Table 17 Clinical features of MS divided by suggestive and not suggestive.**

| Features suggestive of MS: |
|---|
| Relapses and remissions |
| Onset between ages 15 and 50 |
| Optic neuritis |
| Lhermitte's sign |
| Internuclear ophthalmoplegia |
| Fatigue |
| Uhthoff's phenomenon |
| Features not suggestive of MS: |
| Steady progression |
| Onset before age 10 or after age 50 |
| Cortical deficits such as aphasia, apraxia, alexia, neglect |
| Rigidity, sustained dystonia |
| Convulsions |
| Early dementia |
| Deficit developing within minutes |

Not suggestive symptoms and clinical progression descriptions may nonetheless be present in the less common form of MS, primary progressive,

and rare, rapidly fatal (~2 years of onset) malignant forms of MS occur. Table reproduced from (Olek, 2013).

**Table 18 Presenting clinical symptoms in MS divided by gender.**

| Symptom | Females (%) | Males (%) |
|---|---|---|
| Sensory in limbs | 33.2 | 25.1 |
| Visual loss | 16.3 | 15.1 |
| Motor (sub-acute) | 8.3 | 10.4 |
| Diplopia | 6.0 | 8.5 |
| Gait disturbance | 3.2 | 8.3 |
| Motor (acute) | 4.4 | 4.2 |
| Balance problems | 2.5 | 4.0 |
| Sensory in face | 2.9 | 2.5 |
| Lhermitte's sign | 1.6 | 2.3 |
| Vertigo | 1.8 | 1.5 |
| Bladder problems | 0.9 | 1.1 |
| Limb ataxia | 0.9 | 1.3 |
| Acute transverse myelopathy | 0.8 | 0.6 |
| Pain | 0.3 | 0.8 |
| Other | 2.6 | 2.5 |
| Poly-symptomatic onset | 14.5 | 11.9 |

Sensory disturbances in limbs and visual loss, usually temporal with gradual recovery over weeks, full or partial, are the most common. Patients tend to present after a first unrecognised attack and are typically between 20 and 40 years of age. Table reproduced from (Olek, 2013).

**Table 19 Definition of the systemic inflammatory response syndrome (SIRS) and sepsis spectrum in adults.**

| | |
|---|---|
| **SIRS** | **Clinical syndrome resulting from a dysregulated inflammatory response to a non-infectious event** |
| **Sepsis** | **Clinical syndrome resulting from a dysregulated inflammatory response to an infectious event** |
| | **Sepsis criteria** |
| Presence (documented or suspected) of an infection and systemic manifestations including: | |
| General variables | Temperature >38.3 or <36ºC<br>Heart rate >90 bpm or more than two standard deviations above the normal value for age<br>Tachypnea (respiratory rate >20 breaths/minute)<br>Altered mental status<br>Significant edema or positive fluid balance (>20 mL/kg over 24 hours)<br>Hyperglycemia (plasma glucose >140 mg/dL or 7.7 mmol/L) in the absence of diabetes |
| Inflammatory variables | Leukocytosis (WBC count >12,000 microL) or leukopenia (WBC count <4000 microL)<br>Normal WBC count with greater than 10 percent immature forms<br>Plasma C-reactive protein more than two standard deviations above the normal value<br>Plasma procalcitonin more than two standard deviations above the normal value |
| Haemodynamic variables | Arterial hypotension (systolic blood pressure SBP <90 mmHg, MAP <70 mmHg, or an SBP decrease >40 mmHg in adults or less than two standard deviations below normal for age) |
| Organ dysfunction variables | Arterial hypoxemia ($PaO_2/FiO_2$ <300)<br>Acute oliguria (urine output <0.5 mL/kg/hr for at least two hours despite adequate fluid resuscitation)<br>Creatinine increase >0.5 mg/dL or 44.2 micromol/L<br>Coagulation abnormalities (INR >1.5 or aPTT >60 seconds)<br>Ileus (absent bowel sounds)<br>Thrombocytopenia (platelet count <100,000 microL)<br>Hyperbilirubinemia (plasma total bilirubin >4 mg/dL or 70 |

| | |
|---|---|
| | micromol/L) |
| Tissue perfusion variables | Hyperlactatemia (>1 mmol/L)<br>Decreased capillary refill or mottling |
| **Severe sepsis criteria*** | |
| Sepsis-induced tissue hypoperfusion or organ dysfunction with: | |
| Sepsis-induced hypotension<br>Lactate above upper limits of laboratory normal<br>Urine output <0.5 mL/kg/hr for more than two hours despite adequate fluid resuscitation<br>Acute lung injury with $PaO_2/FIO_2$ <250 in the absence of pneumonia as infection source<br>Acute lung injury with $PaO_2/FIO_2$ <200 in the presence of pneumonia as infection source<br>Creatinine >2 mg/dL (176.8 micromol/L)<br>Bilirubin >2 mg/dL (34.2 micromol/L)<br>Platelet count <100,000 microL<br>Coagulopathy (INR >1.5) | |
| Sepsis induced hypotension | SBP <90 mmHg or MAP <70 mmHg or a SBP decrease >40 mmHg or less than two standard deviations below normal for age in the absence of other causes of hypotension. |
| Sepsis-induced tissue hypoperfusion | Infection-induced hypotension, elevated lactate, or oliguria |
| **Septic shock and multiple organ dysfunction syndrome** | |
| Septic shock | Sepsis-induced hypotension persisting despite adequate fluid resuscitation (infusion of 30 mL/kg of crystalloids part of may be albumin equivalent). |
| MODS | Progressive organ dysfunction in an acutely ill patient, such that homeostasis cannot be maintained without intervention.<br>Primary MODS: directly attributable to insult<br>Secondary MODS: organ failure as a consequence of host response |
| Suggested organ specific parameters | $PaO_2/FiO_2$ ratio<br>Platelet count<br>Serum bilirubin<br>Serum creatinine (or urine output)<br>Glasgow coma score<br>Hypotension |

The spectrum and final pathway are thought to be common regardless of the initial insult. SIRS is defined to be initiated by a non-infectious event while sepsis is a response to infection. Both syndromes can derive in MODS and become clinically indistinguishable (Xiao et al., 2011). bpm = beats per minute; WBC = white blood cell; SBP = systolic blood pressure; MAP = mean arterial pressure; INR = international normalized ratio; $PaO_2$ = arterial oxygen tension; $FiO_2$ = fraction of inspired oxygen; aPTT = activated partial thromboplastin time; . *Current discussion argue for including organ failure as a criteria (Vincent et al., 2013). Tables and data reproduced from (Dellinger et al., 2013).

**Table 20 Major mechanisms known to be involved in tissue damage and disease progression during the sepsis and SIRS spectrum.**

| Mechanisms of generalised damage in sepsis | |
|---|---|
| Tissue ischemia | Arterial hypotension and injury to endothelium and microcirculation developed during sepsis worsen tissue perfusion leading to decreased oxygen exchange, tissue ischemia and cellular injury:<br>Microcirculation injury: mediated by alterations in coagulation, fibrinolytic system dysfunction and complement activation.<br>Endothelial damage: likely to be a consequence of exposure to activated PMNs which induce the secretion of ROS, lytic enzymes and vasoactive substances such as NO, PDGF and PAF (Phillipson et al., 2011).<br>Blood cell rigidity: a decrease in PMN and erythrocyte deformability contributes to impaired microvascular flow (Kirschenbaum et al., 2000; Semeraro et al., 2012). |
| Cytopathic injury and cytotoxicity | Thought to be largely a consequence of mitochondrial dysfunction by direct inhibition of respiratory enzyme complexes, oxidative stress damage and mitochondrial DNA breakdown mediated by endotoxin, NO and TNF (Harrois et al., 2009). Release of mitochondrial DAMPs may mimic PAMPs and further induce inflammation (Zhang et al., 2010). Bioenergetic breakdown, as measured by nitric oxide overproduction, antioxidant depletion, mitochondrial dysfunction and decreased ATP concentrations, was associated to organ failure and death in a small cohort of patients with sepsis (Brealey et al., 2002). |
| Apoptosis | Altered apoptosis has been observed in several forms in sepsis studies (Pinheiro da Silva et al., 2009):<br>Pro-inflammatory cytokines may delay apoptosis in activated macrophages and neutrophils;<br>Extensive lymphocyte and dendritic cell apoptosis occurs during sepsis;<br>Lymphocytic apoptosis correlates with severity and immunosuppression;<br>Activation of C5a and other complement members leads to apoptosis and organ dysfunction (Ward, 2008);<br>Inhibition of apoptosis in animals models of sepsis is protective (Coopersmith et al., 2002);<br>Caspasae-3 activation, induction of pro-apoptotic genes (*BIM*, *BID* and *BAK*) and downregulation of anti-apoptotic |

| | |
|---|---|
| | proteins (BCL-2 and BCL-XL) was observed in peripheral blood of sepsis patients (Weber et al., 2008). |
| Immunosuppression | Excess inflammation may be followed by immunosuppression:<br>Autopsy studies reveal a decrease of pro-inflammatory cytokines (TNF, IFNG, IL-6, IL-10) as well as increased expression of inhibitory receptors and ligands and suppressor cell expansion in sepsis deaths as compared to non-sepsis controls (Boomer et al., 2011);<br>Therapy to counter immunosuppression in sepsis may be beneficial (Hotchkiss et al., 2013) |

Systemic effects with cellular responses are underpinned by inflammatory derived damage and tissue hypo-perfusion. PMNs = polymorphonuclear leukocytes; ROS = reactive oxygen species; NO = nitric oxide; PDGF = platelet-derived growth factor; PAF = platelet activating factor; DAMPs = damage associated molecular patterns; PAMPs = pathogen associated molecular patterns; IL = interleukin; TNF = tumour necrosis factor; IFNG = interferon gamma. Data reviewed by (Cohen, 2002; Neviere, 2013) and (Bosmann et al., 2013). Other references are marked within the table.

**Table 21 Major systems and organs affected during the sepsis/SIRS spectrum.**

| Commonly affected organs during sepsis | |
|---|---|
| Circulatory | Hypotension due to vasodilation or redistribution of intravascular fluid is the hallmark of septic shock. Vasodilation is thought to be due to NO and other vasoactive mediators such as prostacyclin. Impaired compensatory vasopressin secretion is thought to be important in vasodilation control in septic patients (Patel et al., 2002; Vincent et al., 2000). |
| Pulmonary | Pulmonary edema can result from endothelial injury in lung vasculature, increased vascular permeability and neutrophil entrapment. Edema leads to hypoxemia with ARDS as a common manifestation in sepsis. |
| Gastrointestinal | Translocation of bacteria due to a breakdown in barrier function may contribute to endotoxin uptake and systemic distribution that leads to MODS (Doig et al., 1998). |
| Hepatic | Two of the most common causes of hepatic dysfunction in sepsis are hypoxic hepatitis and sepsis-associated cholestasis. These lead to dysfunction of the reticuloendothelial system impeding clearing of bacterial and bacterial products originating from the gut. Endotoxin translocation activates Kupffer cells and a pro-inflammatory cascade (Lescot et al., 2012). |
| Renal | Acute renal failure is common in severe sepsis and septic shock (23% and 51% respectively) and increases mortality. The mechanisms are poorly understood but, as in other organs, involve tissue hypo-perfusion and hypoxemia. In the kidney these may lead to acute tubular necrosis and further loss of balance of metabolic and circulatory functions (Schrier et al., 2004). |
| Neurological | Altered CNS manifestations are common and often precede failure in other organs. Encephalopathy is the most common manifestation and thought to be of metabolic origin with inflammatory mediators and blood brain barrier breakdown contributing (Adam et al., 2013). Growing evidence also points to parasympathetic system involvement and acquired paresis (De Jonghe et al., 2002). Chronic neurological damage and cognitive dysfunction following severe sepsis have also been observed (Adam et al., 2013). |

ARDS = acute respiratory distress syndrome; NO = nitric oxide; MODS = multi-organ dysfunction syndrome; CNS = central nervous system. Data reviewed in (Neviere, 2013) and other references within the table.

# Chapter 3: Molecular characterisation of genetic regulatory variants and vitamin D receptor binding in autoimmune disease

## GENETIC VARIATION WITHIN VITAMIN D RECEPTOR BINDING MOTIFS

**Table 22 Enrichment of VDR binding near disease associated traits.**

| Track | Annotation | Observed | Expected | SD | Fold change | P-value | Q-value |
|---|---|---|---|---|---|---|---|
| VDR in LCLs | Hepcidin levels | 17555 | 936 | 1095 | 18.7 | 0.0001 | 0.0009 |
| | Eye color | 28414 | 1820 | 1528 | 15.6 | 0.0001 | 0.0009 |
| | Hematology traits | 77756 | 5209 | 2675 | 14.9 | 0.0001 | 0.0009 |
| | Hair color | 29957 | 2405 | 1642 | 12.5 | 0.0001 | 0.0009 |
| | Haemoglobin | 20897 | 1815 | 1406 | 11.5 | 0.0001 | 0.0009 |
| | Progressive supranuclear palsy | 30347 | 2733 | 1767 | 11.1 | 0.0001 | 0.0009 |
| | Hematocrit | 21486 | 2167 | 1609 | 9.9 | 0.0001 | 0.0009 |
| | Chronic lymphocytic leukemia | 39885 | 4196 | 2302 | 9.5 | 0.0001 | 0.0009 |
| | Coeliac disease and Rheumatoid arthritis | 28523 | 3178 | 1839 | 9.0 | 0.0001 | 0.0009 |
| | Alcohol consumption | 18164 | 2110 | 1562 | 8.6 | 0.0001 | 0.0009 |
| | Tanning | 22217 | 2616 | 1853 | 8.5 | 0.0001 | 0.0009 |
| | Cardiac Troponin-T levels | 27981 | 3591 | 2019 | 7.8 | 0.0001 | 0.0009 |
| | Insulin-like growth factors | 11673 | 1507 | 1286 | 7.7 | 0.0001 | 0.0009 |
| | Mean corpuscular haemoglobin | 22917 | 3653 | 2126 | 6.3 | 0.0001 | 0.0009 |
| | Bilirubin levels | 23978 | 3837 | 2071 | 6.2 | 0.0001 | 0.0009 |
| | Hematological parameters | 23422 | 3967 | 2266 | 5.9 | 0.0001 | 0.0009 |
| | Glycated haemoglobin levels | 19374 | 3409 | 2042 | 5.7 | 0.0001 | 0.0009 |
| | Coeliac disease | 53875 | 9882 | 3391 | 5.5 | 0.0001 | 0.0009 |
| | Asthma | 45266 | 8569 | 3216 | 5.3 | 0.0001 | 0.0009 |
| | Complement C3 and C4 levels | 18418 | 3495 | 2294 | 5.3 | 0.0001 | 0.0009 |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| | Iron status biomarkers | 27985 | 5351 | 2775 | 5.2 | 0.0001 | 0.0009 |
| | IgE levels | 14850 | 2949 | 1914 | 5.0 | 0.0001 | 0.0009 |
| | HIV-1 viral setpoint | 13675 | 2982 | 1879 | 4.6 | 0.0001 | 0.0009 |
| | Graves' disease | 25742 | 5716 | 2974 | 4.5 | 0.0001 | 0.0009 |
| | Mean corpuscular volume | 21816 | 5081 | 2407 | 4.3 | 0.0001 | 0.0009 |
| | Diastolic blood pressure | 26449 | 6314 | 2656 | 4.2 | 0.0001 | 0.0009 |
| | Primary biliary cirrhosis | 28365 | 6841 | 2758 | 4.1 | 0.0001 | 0.0009 |
| | Hypertension | 21054 | 5109 | 2399 | 4.1 | 0.0001 | 0.0009 |
| | Systolic blood pressure | 26109 | 6839 | 2777 | 3.8 | 0.0001 | 0.0009 |
| | Cardiovascular disease risk factors | 21225 | 5900 | 2601 | 3.6 | 0.0001 | 0.0009 |
| | Ankylosing spondylitis | 16529 | 5061 | 2381 | 3.3 | 0.0001 | 0.0009 |
| | Rheumatoid arthritis | 47066 | 14963 | 4151 | 3.1 | 0.0001 | 0.0009 |
| | Orofacial clefts | 30611 | 9852 | 3145 | 3.1 | 0.0001 | 0.0009 |
| | Systemic lupus erythematosus | 46139 | 14973 | 4197 | 3.1 | 0.0001 | 0.0009 |
| | Platelet counts | 50236 | 16934 | 4274 | 3.0 | 0.0001 | 0.0009 |
| | LDL cholesterol | 37553 | 12716 | 3831 | 3.0 | 0.0001 | 0.0009 |
| | Triglycerides | 27609 | 9411 | 3071 | 2.9 | 0.0001 | 0.0009 |
| | Height | 170199 | 65824 | 8386 | 2.6 | 0.0001 | 0.0009 |
| | Cholesterol, total | 36434 | 14674 | 4060 | 2.5 | 0.0001 | 0.0009 |
| | Schizophrenia | 43746 | 19023 | 4565 | 2.3 | 0.0001 | 0.0009 |
| | Blood pressure | 26709 | 11715 | 3541 | 2.3 | 0.0001 | 0.0009 |
| | Ulcerative colitis | 35848 | 15808 | 4183 | 2.3 | 0.0001 | 0.0009 |
| | Crohn's disease | 59961 | 29540 | 5824 | 2.0 | 0.0001 | 0.0009 |
| | Inflammatory bowel disease | 50105 | 27102 | 5508 | 1.8 | 0.0001 | 0.0009 |
| | Obesity-related traits | 217439 | 172002 | 13028 | 1.3 | 0.0001 | 0.0009 |
| | Response to angiotensin II receptor blocker therapy | 18305 | 4638 | 2342 | 3.9 | 0.0003 | 0.0025 |
| | Red blood cell traits | 28281 | 12193 | 3639 | 2.3 | 0.0003 | 0.0025 |
| | Type 1 diabetes | 34116 | 15072 | 4026 | 2.3 | 0.0003 | 0.0025 |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| | Antineutrophil cytoplasmic antibody-associated vasculitis | 10381 | 2243 | 1649 | 4.6 | 0.0007 | 0.0054 |
| | Chronic obstructive pulmonary disease-related biomarkers | 18406 | 5398 | 2651 | 3.4 | 0.0007 | 0.0054 |
| | Visceral adipose tissue/subcutaneous adipose tissue ratio | 24538 | 8675 | 3075 | 2.8 | 0.0007 | 0.0054 |
| | Renal function-related traits (sCR) | 11222 | 2487 | 1665 | 4.5 | 0.0011 | 0.0080 |
| | Atopic dermatitis | 17649 | 5444 | 2532 | 3.2 | 0.0011 | 0.0080 |
| | Smoking behavior | 17270 | 7406 | 2577 | 2.3 | 0.0011 | 0.0080 |
| | Systemic sclerosis | 14099 | 4025 | 2337 | 3.5 | 0.0015 | 0.0107 |
| | Response to Vitamin E supplementation | 9965 | 1822 | 1333 | 5.5 | 0.0023 | 0.0155 |
| | Vertical cup-disc ratio | 9719 | 2415 | 1526 | 4.0 | 0.0023 | 0.0155 |
| | IgA nephropathy | 9496 | 2478 | 1744 | 3.8 | 0.0023 | 0.0155 |
| | Fibrinogen | 7173 | 1806 | 1376 | 4.0 | 0.0027 | 0.0177 |
| | Vitiligo | 21055 | 9557 | 3434 | 2.2 | 0.0035 | 0.0221 |
| | Bipolar disorder and schizophrenia | 40218 | 25029 | 5064 | 1.6 | 0.0039 | 0.0243 |
| | Psoriasis | 17262 | 7164 | 2856 | 2.4 | 0.0043 | 0.0258 |
| | Pulmonary function (interaction) | 23888 | 11103 | 3727 | 2.2 | 0.0043 | 0.0258 |
| | Multiple sclerosis | 44715 | 28841 | 5473 | 1.6 | 0.0047 | 0.0280 |
| | Response to statin therapy | 17361 | 8244 | 3026 | 2.1 | 0.0055 | 0.0315 |
| | Renal function-related traits (eGRFcrea) | 9854 | 3052 | 1996 | 3.2 | 0.0067 | 0.0374 |
| | Cytomegalovirus antibody response | 7655 | 2120 | 1526 | 3.6 | 0.0075 | 0.0412 |
| | Kawasaki disease | 10121 | 3875 | 1962 | 2.6 | 0.0075 | 0.0412 |
| | Hodgkin's lymphoma | 6854 | 1781 | 1488 | 3.8 | 0.0079 | 0.0431 |
| | Cognitive test performance | 0 | 3620 | 2072 | 0.0 | 0.0080 | 0.0435 |
| | Prostate-specific antigen levels | 7560 | 2550 | 1598 | 3.0 | 0.0087 | 0.0470 |

GWAS with eight or more SNPs were tested for higher than expected binding sites of VDR at disease intervals. Traits with significant associations are shown below (n=71) (p-value<0.05, q-value<0.05). A total of 332 traits were tested against VDR binding sites derived from a genome-wide map in lymphoblastoid cells (Ramagopalan et al., 2010) (NHGRI catalogue for 'Multiple sclerosis', http://www.genome.gov/gwastudies/)(Hindorff et al., 2009). Genome-wide binding enrichment of VDR from LCLs was tested using using GAT with the hg19 ungapped contigs as background and 1000 permutations per test to obtain significance values. P-values <0.05 and q-values <0.05 were considered as thresholds. SD = standard deviation; fold = fold change.

**Table 23 Bioinformatics prediction of VDR binding sites containing disease associated SNPs.**

| **>rs7859805 A [A/G] Minor allele count*: A=0.2989/376 (*TRAF1*)** | |
|---|---|
| GTCAAGTCTGACTGGCCGTTGGGGGC[A/G]AAAGGTGTGACTGAAGCACTGACAA | |
| 1 putative site was predicted with binding for the minor [A] allele on the positive strand, 0 sites for the major [G] allele | |
| Score (relative)** | Predicted site sequence |
| 6.754 (0.72) | GGGGCAAAAGGTGTG |
| **>rs1154154 T [C/T] Minor allele count: C=0.2733/329 (*TCRA*)** | |
| ATGGGGAAGTGGGTCCAATGGCCTCT[C/T]TGGCCTCTTAGTACAAAAAGTCTAT | |
| 1 putative site was predicted on the negative strand for the minor allele [T], 0 sites for the major [C] allele | |
| Score (relative)** | Predicted site sequence |
| 7.531 (0.74) | AGGCCAAAGAGGCCA |

*Minor allele counts are derived from a global population (1000Genome phase 1 genotype data) (NCBI dbSNP). *In silico* analysis was carried out using JASPAR CORE vertebrata database (http://jaspar.genereg.net/) with the RXRA::VDR (MA0074.1) Homo sapiens matrix model. **The minimum relative profile score threshold used was 70%.

The tables below present further experimental and predicted evidence of functional impact of disease associated variants found disrupting VDR motifs. The analysis was done using the Regulome database (http://regulome.stanford.edu/).

**Table 24 Scoring system used by Regulome to classify functional evidence of genomic loci.**

| Score | Supporting data |
|---|---|
| 1a | eQTL + TF binding + matched TF motif + matched DNase Footprint + DNase peak |
| 1b | eQTL + TF binding + any motif + DNase Footprint + DNase peak |
| 1c | eQTL + TF binding + matched TF motif + DNase peak |
| 1d | eQTL + TF binding + any motif + DNase peak |
| 1e | eQTL + TF binding + matched TF motif |
| 1f | eQTL + TF binding / DNase peak |
| 2a | TF binding + matched TF motif + matched DNase Footprint + DNase peak |
| 2b | TF binding + any motif + DNase Footprint + DNase peak |
| 2c | TF binding + matched TF motif + DNase peak |
| 3a | TF binding + any motif + DNase peak |
| 3b | TF binding + matched TF motif |
| 4 | TF binding + DNase peak |
| 5 | TF binding or DNase peak |
| 6 | other |

**Table 25 Evidence of functional impact of rs7859805**

| Protein binding over rs7859805 | | | | |
|---|---|---|---|---|
| Method | Location | Bound Protein | Cell Type | Reference |
| ChIP-seq | chr9:123663993-123664249 | POLR2A | GM12878 | ENCODE |
| | chr9:123664015-123664182 | POLR2A | GM12891 | |
| | chr9:123664065-123664166 | POLR2A | GM12891 | |
| | chr9:123663701-123664581 | POLR2A | GM12892 | |
| | chr9:123663701-123664581 | POLR2A | GM12892 | |
| | chr9:123663988-123664276 | POLR2A | GM12892 | |
| | chr9:123663988-123664276 | POLR2A | GM12892 | |
| | chr9:123663976-123664252 | NFKB1 | GM15510 | |
| | chr9:123664096-123664376 | POLR2A | GM18505 | |

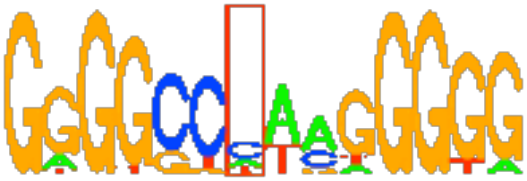
| Motifs covering rs7859805 | | | | | |
|---|---|---|---|---|---|
| Method | Location | Motif | Cell Type | PWM | Reference (PMID) |
| PWM | chr9:123664116-123664130 | PLAG1 | NA | Sequence logo (1) | 18006571 |
| Footprinting | chr9:123664116-123664130 | PLAG1 | GM12892 | Sequence logo (1) | 21106904 |

| Cell types with open chromatin at rs7859805 | | | |
|---|---|---|---|
| Method | Location | Cell Type | Reference |
| DNase-seq | chr9:123664029-123664197 | HEPG2 | ENCODE |
| | chr9:123664080-123664230 | GM12865 | |
| | chr9:123664021-123664367 | CLL | |
| | chr9:123664072-123664333 | GM12892 | |
| | chr9:123664078-123664305 | GM12878 | |
| | chr9:123664086-123664204 | T47D | |

| Gene expression regulation by rs7859805 | | | | | |
|---|---|---|---|---|---|
| Method | Location | Affected Gene | Cell Type | Additional Info | Reference (PMID) |
| eQTL | chr9:123664122-123664123 | PHF19 | Lymphoblastoid | in cis | 20220756 |

**Histone modifications identified over rs7859805 in lymphoblastoid cells***

| Method | Location | Histone Mark | Reference |
|---|---|---|---|
| ChIP-seq | chr9:123620687-123718043 | H3K9me1 | ENCODE |
| | chr9:123620894-123717726 | H3K4me1 | |
| | chr9:123621612-123717063 | H3K27ac | |
| | chr9:123619368-123708457 | H3K79me2 | |
| | chr9:123654044-123697444 | H3K36me3 | |
| | chr9:123644150-123669216 | H3K9me3 | |
| | chr9:123663763-123664198 | H3K4me1 | |
| | chr9:123664020-123664170 | H3K36me3 | |
| | chr9:123648104-123669808 | H3K4me3 | |
| | chr9:123651515-123668921 | H3K27ac | |



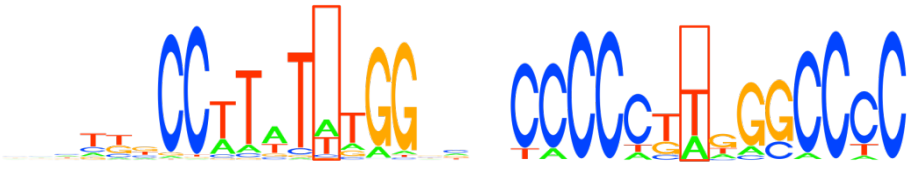
Sequence logo (1)

Chromosomal location of rs7859805 is chr9:123,664,123 (hg19). The score for evidence of functionality was 1b. * GM12878; other cell types (non-lymphoblastoid) showed histone modifications but are not shown for clarity.

**Table 26 Evidence of functional impact of rs1154154**

| Evidence of protein binding over rs1154154 | | | | |
|---|---|---|---|---|
| Method | Location | Bound Protein | Cell Type | Reference |
| ChIP-seq | chr14:23001994-23002337 | POLR2A | GM12878 | ENCODE |
| | chr14:23002010-23002357 | POLR2A | GM12892 | |
| | chr14:23002010-23002357 | POLR2A | GM12892 | |
| | chr14:23001998-23002343 | POLR2A | GM12892 | |
| | chr14:23001998-23002343 | POLR2A | GM12892 | |
| | chr14:23002191-23002441 | BCLAF1 | GM12878 | |
| | chr14:23002181-23002318 | SPI1 | GM12878 | |
| | chr14:23002169-23002312 | SPI1 | GM12891 | |
| | chr14:23002099-23002363 | POLR2A | GM12878 | |
| | chr14:23002153-23002443 | POLR2A | GM12891 | |
| | chr14:23002255-23002515 | POLR2A | GM19099 | |
| **Evidence of motifs covering rs1154154** | | | | |
| Method | Location | Motif | PWM | Reference (PMID) |

| | | | | | |
|---|---|---|---|---|---|
| PWM | chr14:23002259-23002277 | SRF | Sequence logo (2) | 16381825 | |
| | chr14:23002265-23002279 | PLAG1 | Sequence logo (3) | 18006571 | |
| **Cell types with open chromatin at rs1154154** | | | | | |
| Method | Location | | Cell Type | | Reference |
| DNase-seq | chr14:23002180-23002330 | | GM12865 | | ENCODE |
| FAIRE | chr14:23001989-23002752 | | GM12892 | | |
| | chr14:23002129-23003350 | | GM18507 | | |
| | chr14:22999478-23002683 | | GM19239 | | |
| **Histone modifications identified over rs1154154 in lymphoblastoid cells** | | | | | |
| Method | | Location | | Histone Mark | Reference |
| ChIP-seq | | chr14:21922101-23025267 | | H3K27me3 | ENCODE |
| | | chr14:22838448-23077802 | | H2.AZ | |
| | | chr14:22871955-23091570 | | H3K4me1 | |
| | | chr14:22872021-23089994 | | H3K4me1 | |
| | | chr14:22877685-23078959 | | H3K04me3 | |
| | | chr14:22878021-23079296 | | H3K4me2 | |
| | | chr14:22907395-23077908 | | H3K27ac | |
| | | chr14:22907772-23078582 | | H3K9ac | |
| | | chr14:22908213-23078695 | | H3K4me3 | |
| | | chr14:22925800-23078022 | | H3K79me2 | |
| | | chr14:22926473-23088908 | | H3K36me3 | |
| | | chr14:22973049-23017627 | | H3K9me3 | |



Sequence logo (2) and (3)

Chromosomal location of rs1154154 is chr14:23,002,272 (hg19). The score for evidence of functionality was 3a. * GM12878; other cell types (non-lymphoblastoid) showed histone modifications but are not shown for clarity.

# Chapter 4: Integrative analysis of vitamin D in healthy individuals

**Table 27 Serum 25(OH)$D_3$ in healthy volunteers according to gender**

| 25(OH)$D_3$ | Sex | | Total |
|---|---|---|---|
| | Male | Female | |
| **Valid N** | 112 | 145 | 257 |
| **Mean** | 54.6 | 64.4 | 60.1 |
| **Standard Deviation** | 20.1 | 24.9 | 23.4 |
| **Median** | 52.3 | 60.7 | 58.5 |
| **Percentile 25** | 40 | 45.6 | 43.6 |
| **Percentile 75** | 65.3 | 80 | 74 |
| **CI 95%** | 60.7-68 | 50.5-58.6 | 57.2-62.9 |

**Table 28 Serum levels of 25(OH)$D_3$ in healthy subjects by month and gender**

| 25(OH)$D_3$ | Date by month | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | Mar | | Apr | | May | | Jun | | Jul | |
| | Sex | | | | | | | | | |
| | M | F | M | F | M | F | M | F | M | F |
| **Count** | 0 | 8 | 28 | 31 | 15 | 13 | 49 | 47 | 23 | 47 |
| **Mean** | . | 40 | 41 | 50 | 59 | 64 | 56 | 70 | 64 | 73 |
| **Standard Deviation** | . | 28 | 17 | 21 | 27 | 28 | 15 | 21 | 21 | 24 |
| **Minimum** | . | 17 | 12 | 16 | 17 | 9 | 34 | 31 | 35 | 15 |
| **Maximum** | . | 90 | 74 | 111 | 108 | 102 | 89 | 157 | 106 | 129 |

M = male; F = female

**Table 29 Expressed monocyte genes correlated to plasma vitamin D in healthy volunteers**

| ProbeID | TargetID | PROBE_ID | CHROMOSOME | PROBE_COORDINATES | FDR | P | R |
|---|---|---|---|---|---|---|---|
| 7050561 | ZNF259 | ILMN_1753790 | 11 | 116649657-116649706 | 0.010 | 4.29E-07 | -0.308 |
| 3460056 | C1ORF57 | ILMN_1657446 | 1 | 231180561-231180600:231180601-231180610 | 0.010 | 6.77E-07 | -0.303 |
| 650300 | CCDC86 | ILMN_1792681 | 11 | 60374905-60374954 | 0.010 | 8.18E-07 | -0.301 |
| 2320292 | TAF5L | ILMN_2252813 | 1 | 229729411-229729460 | 0.010 | 1.12E-06 | -0.297 |
| 6760504 | SEH1L | ILMN_2381138 | 18 | 12955607-12955608:12961152-12961199 | 0.010 | 1.41E-06 | -0.294 |
| 1470376 | SCYL1 | ILMN_1731991 | 11 | 65062533-65062582 | 0.010 | 1.45E-06 | -0.294 |
| 2970482 | RNU1-5 | ILMN_3236653 | 1 | 149224093-149224142 | 0.010 | 1.54E-06 | 0.293 |
| 7560092 | TMEM126A | ILMN_2217809 | 11 | 85042843-85042892 | 0.011 | 1.83E-06 | -0.291 |
| 2970296 | ALG8 | ILMN_2365686 | 11 | 77815465-77815499:77817853-77817867 | 0.015 | 2.80E-06 | -0.286 |
| 5080546 | MRPL12 | ILMN_1699603 | 17 | 77284452-77284475:77284559-77284584 | 0.015 | 3.15E-06 | -0.285 |
| 2940228 | HS.36053 | ILMN_1819854 | 5 | 111088701-111088750 | 0.017 | 3.85E-06 | 0.283 |
| 2140176 | THYN1 | ILMN_2357361 | 11 | 134119131-134119156:134119722-134119745 | 0.019 | 4.74E-06 | -0.280 |
| 2320093 | PIGY | ILMN_1672366 | 4 | 89661596-89661645 | 0.021 | 7.00E-06 | -0.275 |
| 730092 | SFXN4 | ILMN_2363361 | 10 | 120895771-120895820 | 0.021 | 8.17E-06 | -0.273 |
| 4010343 | TYW3 | ILMN_1726842 | 1 | 75003783-75003832 | 0.021 | 1.05E-05 | -0.270 |
| 1400341 | ACY1 | ILMN_1683883 | 3 | 51997630-51997659:51997822-51997841 | 0.021 | 1.13E-05 | -0.269 |
| 4280136 | ADRM1 | ILMN_2389013 | 20 | 60316831-60316880 | 0.021 | 1.13E-05 | -0.269 |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| 1820189 | ICT1 | ILMN_2182198 | 17 | 70528667-70528695:70528696-70528716 | 0.021 | 1.18E-05 | -0.269 |
| 7560253 | EXOSC4 | ILMN_1745271 | 8 | 145135049-145135052:145135145-145135190 | 0.021 | 1.20E-05 | -0.268 |
| 6650747 | ZNF641 | ILMN_1731666 | 12 | 48736083-48736132 | 0.021 | 1.25E-05 | -0.268 |
| 1580427 | THEM2 | ILMN_2098743 | 6 | 24806143-24806192 | 0.021 | 1.38E-05 | -0.267 |
| 160114 | SMAP1 | ILMN_1768271 | 6 | 71628034-71628083 | 0.021 | 1.42E-05 | -0.266 |
| 6660047 | KIAA0090 | ILMN_2153280 | 1 | 19544603-19544652 | 0.021 | 1.44E-05 | -0.266 |
| 5270280 | PA2G4 | ILMN_1728984 | 12 | 54793008-54793057 | 0.021 | 1.56E-05 | -0.265 |
| 840706 | PRICKLE3 | ILMN_1656942 | X | 49032066-49032115 | 0.021 | 1.71E-05 | -0.264 |
| 5390561 | UNC50 | ILMN_1766981 | 2 | 98601096-98601145 | 0.021 | 1.75E-05 | -0.263 |
| 2120204 | AGPAT3 | ILMN_2288175 | 21 | 44148278-44148327 | 0.021 | 1.85E-05 | -0.263 |
| 3870041 | KCNK6 | ILMN_2074773 | 19 | 38819182-38819231 | 0.021 | 1.91E-05 | -0.262 |
| 6760730 | RSRC1 | ILMN_1682494 | 3 | 158261224-158261273 | 0.021 | 1.96E-05 | -0.262 |
| 4860356 | BYSL | ILMN_1682792 | 6 | 41900533-41900582 | 0.021 | 1.96E-05 | -0.262 |
| 6370519 | HGSNAT | ILMN_1662413 | 8 | 43053022-43053071 | 0.021 | 2.00E-05 | -0.262 |
| 5050575 | ASB13 | ILMN_1654385 | 10 | 5721023-5721072 | 0.021 | 2.02E-05 | -0.262 |
| 630204 | RAC1 | ILMN_2359789 | 7 | 6409822-6409871 | 0.021 | 2.03E-05 | -0.261 |
| 1050367 | GGA3 | ILMN_2402972 | 17 | 73233242-73233291 | 0.021 | 2.17E-05 | -0.261 |
| 5090035 | PPP2R5D | ILMN_1780940 | 6 | 42979773-42979822 | 0.021 | 2.18E-05 | -0.260 |
| 2070403 | CABLES2 | ILMN_1762407 | 20 | 60963850-60963899 | 0.021 | 2.22E-05 | -0.260 |
| 6650017 | ATP6V1B2 | ILMN_1787705 | 8 | 20078945-20078994 | 0.021 | 2.24E-05 | -0.260 |
| 2480241 | ATP6V1C1 | ILMN_1659801 | 8 | 104085004-104085053 | 0.021 | 2.29E-05 | -0.260 |
| 5570091 | RANBP3 | ILMN_1723689 | 19 | 5867384-5867433 | 0.021 | 2.30E-05 | -0.260 |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| 1090687 | POLR1D | ILMN_1742427 | 13 | 27139220-27139269 | 0.021 | 2.42E-05 | -0.259 |
| 730605 | MRPS22 | ILMN_1655377 | 3 | 139074572-139074621 | 0.021 | 2.45E-05 | -0.259 |
| 1410161 | KLHL5 | ILMN_1706687 | 4 | 38799958-38800007 | 0.021 | 2.56E-05 | -0.258 |
| 5960703 | Sep-02 | ILMN_1748546 | 2 | 241941730-241941779 | 0.021 | 2.65E-05 | -0.258 |
| 2470427 | PSME3 | ILMN_1800975 | 17 | 38249065-38249114 | 0.021 | 2.66E-05 | -0.258 |
| 2570441 | TFIP11 | ILMN_2408102 | 22 | 26888084-26888133 | 0.021 | 2.69E-05 | -0.258 |
| 3360220 | RMND1 | ILMN_1694533 | 6 | 151768567-151768616 | 0.021 | 2.74E-05 | -0.257 |
| 2140528 | GPR172A | ILMN_2041577 | 8 | 145555566-145555615 | 0.021 | 2.80E-05 | -0.257 |
| 5570279 | HIST1H1C | ILMN_1757406 | 6 | 26164078-26164127 | 0.021 | 2.76E-05 | 0.257 |
| 840356 | NUAK2 | ILMN_2094952 | 1 | 205271519-205271568 | 0.021 | 2.63E-05 | 0.258 |
| 7210035 | SNORD13 | ILMN_1892403 | 8 | 33371010-33371059 | 0.021 | 2.46E-05 | 0.259 |
| 5670152 | RNU1-3 | ILMN_3246273 | 1 | 16993310-16993359 | 0.021 | 2.43E-05 | 0.259 |
| 1820296 | WTAP | ILMN_2279339 | 6 | 160096993-160097042 | 0.021 | 2.35E-05 | 0.260 |
| 1980201 | WWP1 | ILMN_1804328 | 8 | 87479024-87479073 | 0.021 | 2.33E-05 | 0.260 |
| 3850189 | RNU4-2 | ILMN_3308138 | | | 0.021 | 2.24E-05 | 0.260 |
| 2260349 | MIR1974 | ILMN_3308961 | | | 0.021 | 2.18E-05 | 0.260 |
| 5080066 | BRWD1 | ILMN_1673518 | 21 | 40668602-40668651 | 0.021 | 2.09E-05 | 0.261 |
| 4070608 | CBLL1 | ILMN_2073732 | 7 | 107187084-107187133 | 0.021 | 1.83E-05 | 0.263 |
| 3610040 | C9ORF93 | ILMN_1764188 | 9 | 15724816-15724865 | 0.021 | 1.67E-05 | 0.264 |
| 6510367 | JUN | ILMN_1806023 | 1 | 59247250-59247299 | 0.021 | 1.38E-05 | 0.267 |
| 7550315 | C17ORF108 | ILMN_2099487 | 17 | 26205699-26205748 | 0.021 | 1.04E-05 | 0.270 |
| 5690113 | NUFIP2 | ILMN_2102693 | 17 | 24612861-24612910 | 0.021 | 8.32E-06 | 0.273 |
| 5960086 | HS.562219 | ILMN_1901419 | 1 | 28847792-28847841 | 0.021 | 6.63E-06 | 0.276 |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| 2650091 | LOC648390 | ILMN_3287583 | 2 | 136803801-136803850 | 0.021 | 6.30E-06 | 0.276 |
| 5490411 | CCDC22 | ILMN_1781234 | X | 48993563-48993612 | 0.021 | 2.85E-05 | -0.257 |
| 650022 | USP33 | ILMN_2306077 | 1 | 78191317-78191366 | 0.021 | 2.92E-05 | -0.257 |
| 4230259 | BRCC3 | ILMN_1697546 | X | 154003954-154004003 | 0.021 | 2.94E-05 | -0.256 |
| 6550609 | MYO1A | ILMN_1704861 | 12 | 57422486-57422535 | 0.022 | 3.17E-05 | -0.255 |
| 5720538 | IPO7 | ILMN_1652378 | 11 | 9423259-9423308 | 0.023 | 3.35E-05 | -0.255 |
| 2140228 | CDR2 | ILMN_1720270 | 16 | 22357429-22357478 | 0.023 | 3.34E-05 | 0.255 |
| 4890382 | ILVBL | ILMN_1759419 | 19 | 15226152-15226201 | 0.023 | 3.41E-05 | -0.254 |
| 2260176 | SRFBP1 | ILMN_1721024 | 5 | 121390610-121390659 | 0.023 | 3.49E-05 | -0.254 |
| 5260328 | HS.403972 | ILMN_1864166 | 1 | 41745269-41745318 | 0.023 | 3.60E-05 | -0.254 |
| 7650468 | TMEM168 | ILMN_1758679 | 7 | 112407428-112407477 | 0.023 | 3.64E-05 | -0.254 |
| 1980239 | LOC649553 | ILMN_3200830 | 7 | 64168295-64168297:64168722-64168768 | 0.023 | 3.74E-05 | -0.253 |
| 3460220 | VPS4A | ILMN_1708946 | 16 | 67916192-67916241 | 0.023 | 3.78E-05 | -0.253 |
| 1300041 | C19ORF42 | ILMN_1694759 | 19 | 16757164-16757213 | 0.023 | 3.83E-05 | -0.253 |
| 2060121 | FUCA1 | ILMN_1752728 | 1 | 24044404-24044453 | 0.023 | 3.86E-05 | -0.253 |
| 6220112 | SCARNA9 | ILMN_1805064 | 11 | 93094373-93094422 | 0.023 | 3.82E-05 | 0.253 |
| 3440301 | LIN54 | ILMN_1724062 | 4 | 84065267-84065316 | 0.023 | 3.64E-05 | 0.254 |
| 4760156 | LOC100134359 | ILMN_3246519 | | | 0.024 | 4.00E-05 | 0.252 |
| 7330630 | BTF3 | ILMN_1659762 | 5 | 72834636-72834685 | 0.024 | 4.07E-05 | -0.252 |
| 7320717 | GALNT4 | ILMN_1739297 | 12 | 89915490-89915539 | 0.024 | 4.21E-05 | -0.252 |
| 7160440 | ARID4A | ILMN_1810229 | 14 | 57908957-57909006 | 0.025 | 4.34E-05 | 0.251 |
| 6100441 | PRPF4 | ILMN_1697440 | 9 | 115094549-115094598 | 0.025 | 4.49E-05 | -0.251 |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| 1740040 | CLN5 | ILMN_1778203 | 13 | 76474488-76474537 | 0.026 | 4.66E-05 | -0.250 |
| 3420377 | RPRD2 | ILMN_3238889 | 1 | 150448791-150448840 | 0.026 | 4.65E-05 | 0.250 |
| 7650725 | UGCGL1 | ILMN_2360291 | 2 | 128665348-128665397 | 0.026 | 4.73E-05 | -0.250 |
| 4220576 | SLC4A1AP | ILMN_1750876 | 2 | 27770111-27770125:27771005-27771039 | 0.026 | 4.96E-05 | -0.249 |
| 3440368 | PPP1R7 | ILMN_1808333 | 2 | 241757953-241757955:241770735-241770781 | 0.026 | 5.17E-05 | -0.249 |
| 6290356 | SERPINF1 | ILMN_2141482 | 17 | 1627514-1627563 | 0.026 | 5.18E-05 | -0.249 |
| 1710102 | EIF2C1 | ILMN_1671326 | 1 | 36162240-36162289 | 0.026 | 5.32E-05 | -0.248 |
| 3780168 | COL8A2 | ILMN_1674050 | 1 | 36561367-36561416 | 0.026 | 5.41E-05 | -0.248 |
| 630446 | C22ORF25 | ILMN_1789405 | 22 | 18433229-18433278 | 0.026 | 5.48E-05 | -0.248 |
| 3120750 | SEMA6C | ILMN_1813503 | 1 | 151104333-151104382 | 0.026 | 5.49E-05 | -0.248 |
| 60070 | LPGAT1 | ILMN_2151277 | 1 | 209986819-209986868 | 0.026 | 5.55E-05 | -0.248 |
| 3370687 | CGGBP1 | ILMN_1752631 | 3 | 88184245-88184294 | 0.026 | 5.56E-05 | 0.248 |
| 5890176 | LOC442454 | ILMN_1711729 | X | 56780689-56780738 | 0.026 | 5.41E-05 | 0.248 |
| 160079 | MGC10997 | ILMN_2147251 | 15 | 38092941-38092990 | 0.026 | 5.28E-05 | 0.248 |
| 2120114 | DYRK1A | ILMN_2374293 | 21 | 37808642-37808691 | 0.026 | 5.10E-05 | 0.249 |
| 4060692 | SUCLG2 | ILMN_1652379 | 3 | 67507966-67508015 | 0.026 | 4.95E-05 | 0.249 |

Top 100 genes of 587 (FDR <0.05) shown ranked by significance.

# Chapter 5: Clinical and molecular correlates of vitamin D and severe sepsis survival

## CLINICAL AND DEMOGRAPHIC CHARACTERISTICS OF THE GAINS COHORT

**Table 30 Clinical and demographic characteristics of the GAinS cohort (numerical variables run in regression analysis and others of importance).**

| Variable | Valid N | Missing | Median | Percentile 25 | Percentile 75 |
|---|---|---|---|---|---|
| Age (years) | 143 | 0 | 64 | 52 | 76 |
| APACHE II score | 142 | 1 | 17.5 | 14 | 22 |
| Bilirubin (mmol/l) | 139 | 4 | 11 | 7 | 19 |
| Charlson score | 143 | 0 | 0 | 0 | 2 |
| CO2 arterial pressure (kPa) | 141 | 2 | 5.1 | 4.5 | 6.1 |
| Days in ICU | 143 | 0 | 5 | 3 | 14 |
| Mean arterial pressure (mmHg) | 143 | 0 | 63 | 55 | 68 |
| O2 arterial pressure (kPa) | 141 | 2 | 9.14 | 7.9 | 11 |
| Plasma 25(OH)$D_3$ (nmol/L) | 143 | 0 | 14 | 9 | 25 |
| Plasma cathelicidin (ng/ml) | 139 | 4 | 35.84 | 31.63 | 39.46 |
| Sodium (mmol/l) | 142 | 1 | 135.5 | 133 | 138 |
| SOFA (highest) | 142 | 1 | 8 | 5 | 10 |
| Lymphocytes (*10E9/l) | 134 | 9 | 0.79 | 0.46 | 1.16 |
| Monocytes (*10E9/l) | 133 | 10 | 0.56 | 0.26 | 0.84 |
| Polynucleocytes (*10E9/l) | 135 | 8 | 10 | 6.9 | 16.64 |
| Respiratory rate (bpm) | 141 | 2 | 26 | 16 | 34 |

**Table 31 Clinical and demographic characteristics of the GAinS cohort (categorical variables used in the regression analysis and others of importance).**

| Variable | | Count | Column N % |
|---|---|---|---|
| Diagnosis | CAP | 77 | 53.8 |
| | FP | 66 | 46.2 |
| Ethnicity | African | 2 | 1.4 |
| | Asian | 6 | 4.2 |
| | Caucasian | 133 | 93.0 |
| | Mediterranean | 2 | 1.4 |
| Season of admission | Autumn | 35 | 24.5 |
| | Spring | 34 | 23.8 |
| | Summer | 33 | 23.1 |
| | Winter | 41 | 28.7 |
| Vasopressors/Inotrope therapy* | High | 69 | 48.3 |
| | Moderate | 13 | 9.1 |
| | Low | 1 | 0.7 |
| | None | 60 | 42.0 |
| Renal replacement therapy | FALSE | 132 | 92.3 |
| | TRUE | 11 | 7.7 |
| Gender | Female | 70 | 49.0 |
| | Male | 73 | 51.0 |

*Defined as following: None = No hypotension; Low = Dopamine <= 5 $\mu$g/kg/min or Dobutamine; Moderate = Dopamine > 5 $\mu$g/kg/min, epi or norepi <=0.1 $\mu$g/kg/min; High = Dopamine > 15 $\mu$g/kg/min, epi or norepi >0.1 $\mu$g/kg/min.

**Table 32 25(OH)$D_3$ serum levels in patients with severe sepsis or septic shock according to survival and day of sampling**

| | Survival at ICU discharge | | | | | | | | Survival at 28 days | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Alive | | | | Dead | | | | Alive | | | | Dead | | | |
| | N | Median | P 25 | P 75 | N | Median | P 25 | P 75 | N | Median | P 25 | P 75 | N | Median | P 25 | P 75 |
| **1st day** | 68 | 12 | 9 | 21 | 17 | 11 | 9 | 14 | 61 | 12 | 9 | 21 | 23 | 11 | 9 | 15 |
| **3rd day** | 75 | 18 | 10 | 27 | 10 | 14 | 13 | 15 | 71 | 18 | 10 | 27 | 14 | 15 | 14 | 20 |
| **5th day** | 64 | 15 | 12 | 26 | 8 | 10 | 9 | 23 | 64 | 15 | 12 | 28 | 7 | 17 | 9 | 25 |
| **1st available** | 121 | 14 | 9 | 25 | 23 | 11 | 9 | 15 | 114 | 14 | 9 | 25 | 28 | 11 | 9 | 16 |
| | Survival at hospital discharge | | | | | | | | Survival at 6 months | | | | | | | |
| | Alive | | | | Dead | | | | Alive | | | | Dead | | | |
| | N | Median | P 25 | P 75 | N | Median | P 25 | P 75 | N | Median | P 25 | P 75 | N | Median | P 25 | P 75 |
| **1st day** | 59 | 13 | 9 | 21 | 26 | 10 | 9 | 14 | 54 | 13 | 9 | 21 | 30 | 10 | 9 | 14 |
| **3rd day** | 66 | 18 | 11 | 27 | 19 | 14 | 10 | 16 | 61 | 20 | 11 | 27 | 24 | 14 | 11 | 16 |
| **5th day** | 58 | 16 | 13 | 30 | 14 | 10 | 9 | 18 | 55 | 17 | 13 | 32 | 16 | 12 | 9 | 18 |
| **1st available** | 109 | 16 | 10 | 26 | 35 | 10 | 9 | 15 | 100 | 16 | 10 | 26 | 42 | 11 | 9 | 15 |

**PAIR-WISE COMPARISONS**

**Table 33 Serial sampling of 25(OH)$D_3$ plasma levels in patients with severe sepsis or septic shock**

| Plasma 25(OH)$D_3$, n=32 | n | Median | Percentile 25 | Percentile 75 |
|---|---|---|---|---|
| Day 1 | 32 | 12.5 | 9.0 | 19.5 |
| Day 3 | 32 | 15.0 | 11.0 | 21.0 |
| Day 5 | 32 | 15.0 | 12.0 | 24.0 |
| First available sample | 143 | 14.0 | 9.0 | 25.0 |

Values for days 1, 3 and 5 represent patients for whom all samples were available. The first available sample was used for the full cohort and subsequent analyses.

**Table 34 Significance test values for 25(OH)$D_3$ plasma levels and survival**

| 25(OH)$D_3$ | Survival | | | |
|---|---|---|---|---|
| Type of sample | ICU discharge | 28 days | Hospital discharge | 6 months |
| Day 1 | 0.320 | 0.526 | 0.058 | 0.016 |
| Day 3 | 0.437 | 0.810 | 0.092 | 0.040 |
| Day 5 | 0.157 | 0.381 | 0.028 | 0.021 |
| 1st available sample | 0.166 | 0.121 | 0.011 | 0.003 |

Independent samples Mann Whitney U test, exact two tailed p values (day 1 n=84, day 3 n=40, day 5 n=19).

**Table 35 Test of proportions according to survival and serum 25(OH)$D_3$**

| 25(OH)$D_3$ (VD3_day_median) | | Survival at ICU discharge | | Survival at 28 days | | Survival at hospital discharge | | Survival at 6 months | |
|---|---|---|---|---|---|---|---|---|---|
| | | **Alive** | **Dead** | **Alive** | **Dead** | **Alive** | **Dead** | **Alive** | **Dead** |
| | | Column N % | | | | | | | |
| **VD3_1_13** | ≤median | 57.4%$_a$ | 64.7%$_a$ | 59.0%$_a$ | 60.9%$_a$ | 54.2%$_a$ | 69.2%$_a$ | 53.7%$_a$ | 70.0%$_a$ |
| | >median | 42.6%$_a$ | 35.3%$_a$ | 41.0%$_a$ | 39.1%$_a$ | 45.8%$_a$ | 30.8%$_a$ | 46.3%$_a$ | 30.0%$_a$ |
| **VD3_3_18** | ≤median | 54.7%$_a$ | 80.0%$_a$ | 54.9%$_a$ | 71.4%$_a$ | 51.5%$_a$ | 78.9%$_b$ | 47.5%$_a$ | 83.3%$_b$ |
| | >median | 45.3%$_a$ | 20.0%$_a$ | 45.1%$_a$ | 28.6%$_a$ | 48.5%$_a$ | 21.1%$_b$ | 52.5%$_a$ | 16.7%$_b$ |
| **VD3_5_16** | ≤median | 51.6%$_a$ | 62.5%$_a$ | 53.1%$_a$ | 42.9%$_a$ | 50.0%$_a$ | 64.3%$_a$ | 49.1%$_a$ | 62.5%$_a$ |
| | >median | 48.4%$_a$ | 37.5%$_a$ | 46.9%$_a$ | 57.1%$_a$ | 50.0%$_a$ | 35.7%$_a$ | 50.9%$_a$ | 37.5%$_a$ |
| **VD3_1ST_16** | ≤median | 54.5%$_a$ | 78.3%$_b$ | 54.4%$_a$ | 75.0%$_b$ | 52.3%$_a$ | 77.1%$_b$ | 50.0%$_a$ | 78.6%$_b$ |
| | >median | 45.5%$_a$ | 21.7%$_b$ | 45.6%$_a$ | 25.0%$_b$ | 47.7%$_a$ | 22.9%$_b$ | 50.0%$_a$ | 21.4%$_b$ |

Highlighted values are significantly different at $p< 0.05$, two-sided test of equality for column proportions assuming equal variances. Cells with no subscript are not included.

**Table 36 Pearson chi square tests of independence for survival and 25(OH)$D_3$ plasma levels**

| Sample | Survival at ICU discharge | Survival at 28 days | Survival at hospital discharge | Survival at 6 months |
|---|---|---|---|---|
| **VD3_1_13** | 0.582 | 0.877 | 0.196 | 0.145 |
| **VD3_3_18** | ---- | 0.254 | .033* | .003* |
| **VD3_5_16** | ---- | ---- | 0.337 | 0.345 |
| **VD3_1ST_16** | .034* | .047* | .009* | .002* |

Values not shown had insufficient cell counts. Labels = VD3_day of sampling_median(nmol/L); 1st means the first available sample (either day 1, 3 or 5).

**REGRESSION ANALYSIS**

**Table 37 Variables used in regression model to predict survival**

| Variable | Coding |
|---|---|
| 25(OH)$D_3$ | Levels in blood on first day or closest available to ICU admission |
| Age | Age in years |
| APACHE II | Score at ICU admission |
| Bicarbonate | Levels in blood on first day or closest available to ICU admission |
| Bilirubin | Levels in blood on first day or closest available to ICU admission |
| Blood pH | Levels in blood on first day or closest available to ICU admission |
| Blood urea nitrogen | Levels in blood on first day or closest available to ICU admission |
| Cathelicidin (LL-37) | Cathelicidin levels in blood (same day as 25(OH)$D_3$) |
| Charlson score | Updated Charlson index (1) |

| | |
|---|---|
| CO2 arterial pressure | mmHg |
| Creatinin | Highest value/lowest value |
| Days in ICU | Number of days in ICU |
| Diagnosis | CAP/FP |
| Ethnicity | |
| Fraction of inspired oxygen | Percentage |
| Heart rate | Highest value/lowest value |
| Hematocrit | Count on day one of admission |
| Hospital | ICU facility |
| Ionotrope therapy | Low, moderate, high or none |
| Lymphocytes | Count on day one of admission |
| Mean arterial pressure | Highest value/lowest value |
| Monocytes | Count on day one of admission |
| O2 arterial pressure | mmHg |
| Platelet count | Levels in blood on first day or closest available to ICU admission |
| Polynucleocytes | Count on day one of admission |
| Potassium | Highest value/lowest value |
| Renal failure | Presence/Absence |
| Renal replacement therapy | Presence/Absence |
| Respiratory rate | Highest value/lowest value |
| Season of admission | Spring, summer, autumn, winter |
| Sex | Female/Male |
| Sodium | Highest value/lowest value |
| SOFA | Highest SOFA score in 7 day period (or less) |
| Systolic blood pressure | Highest value/lowest value |
| Temperature | Highest value/lowest value |
| Type of operation | Emergency/None |
| Urinary volume | 24 hour period |
| Use of ventilator | True/False |

Highest/lowest variables (=two variables) correspond to recordings on day one of ICU admission or first available thereof. Variables that were significantly correlated between each other were excluded from subsequent regression analysis (supplementary figures below). Outlier analysis was performed for each model (supplementary tables ). (1) I used Quan et al 2011 (Quan et al., 2011) with modifications: for rheumaotologic disease we also included systemic lupus erythematousus.

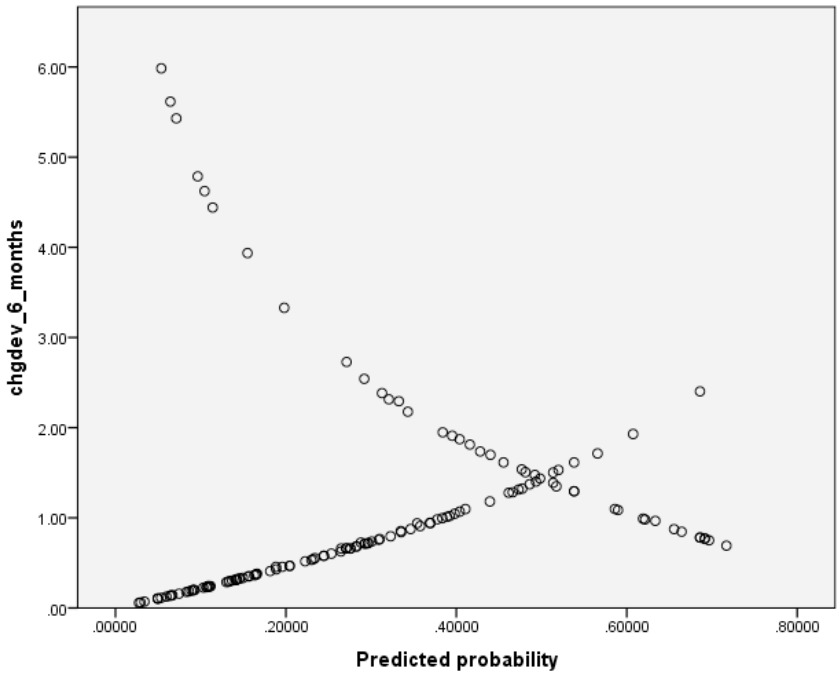



**Figure 31 Outlier analysis: change in deviation of standard residuals and predicted probability.**

Outliers were analysed using a change in deviation plot to identify cases that fit the model poorly. I identified six possible outliers with changes in deviation (>4) using survival at six months in the forward conditional logistic

regression model. We performed the same analysis for forward and backward models at each survival time-point (data not shown).

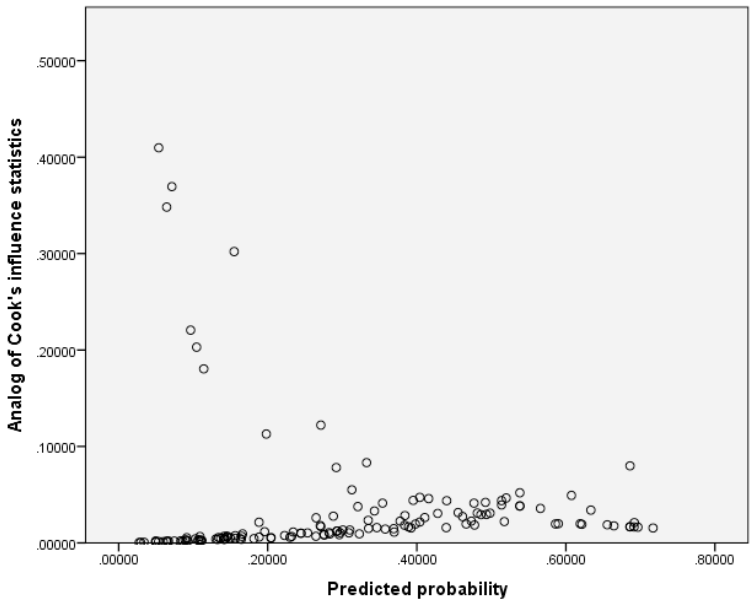



**Figure 32 Outlier analysis: Influential cases.**

I used Cook's distance to identify cases that appeared overly influential to the model according to the predicted probability grouped by survival at six months. The six cases identified by high values in deviation changes also had high values for Cook's distance and were excluded. We performed the same analysis for forward and backward models at each survival time-point (data not shown).

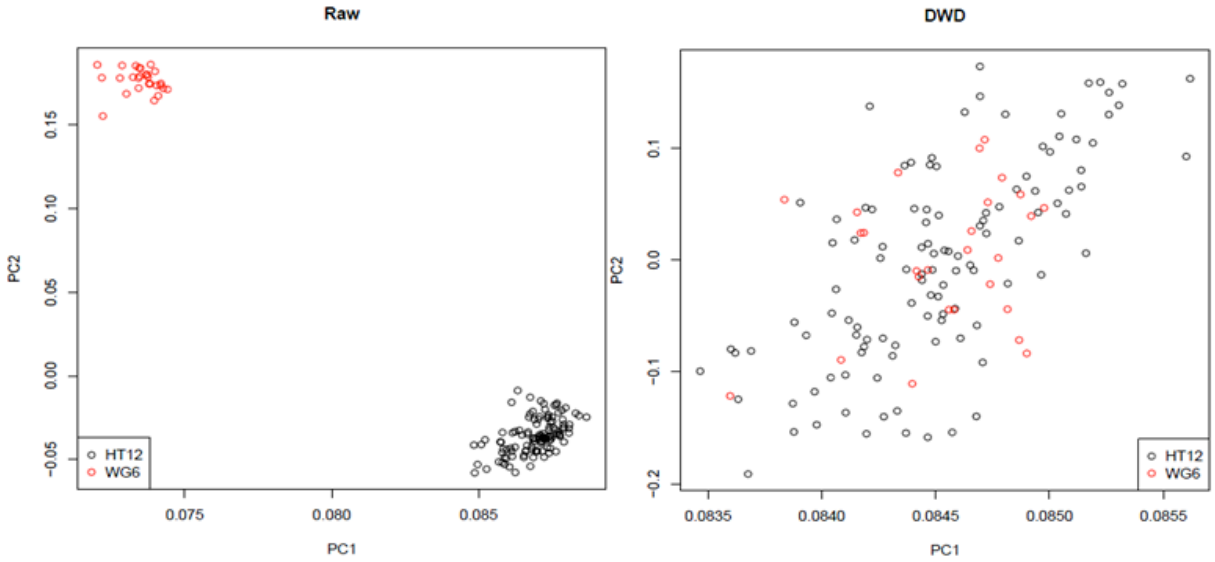



**Figure 33 Principal component analysis of gene expression data in the GAinS cohort.**

Two platforms had been used originally as explained in the methods section. In this study they were normalised using distance weighted discrimination (DWD) method as it is robust for different sizes and different platforms (Rudy et al., 2011). The plot on the left shows the raw data and the one on the right after normalisation.

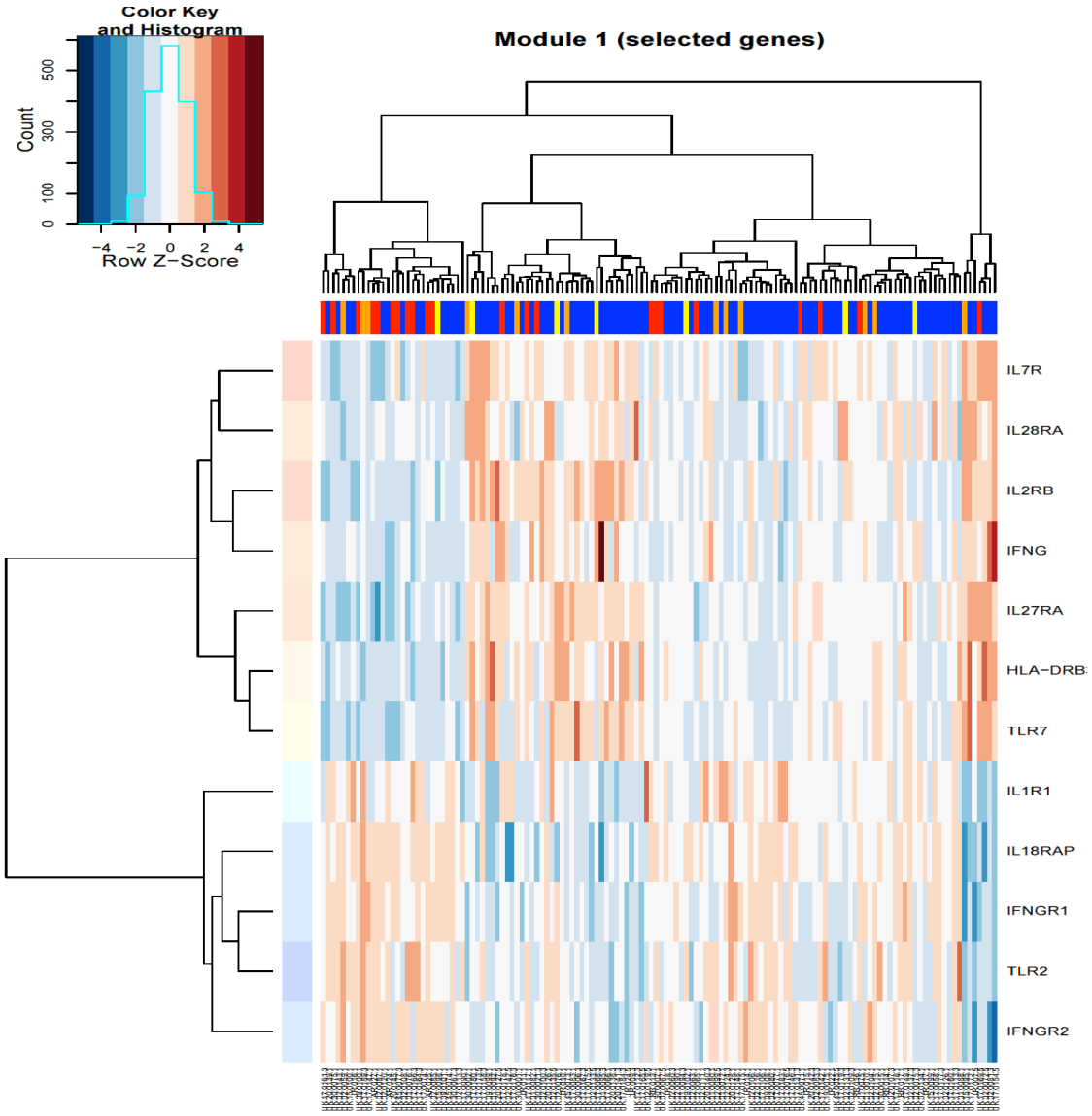



**Figure 34 Expressed leukocyte genes from patients with severe sepsis and septic shock correlated to vitamin D.**

As heatmap in the main text for genes of interest: hierarchical clustering of genes expressed in patients with severe sepsis and septic shock and correlated to plasma 25(OH)D (FDR 5%). Results on the vertical axis were arranged according to correlation with 25(OH)D (red, positive; blue,

negative, Spearman rho). Gene expression values represent Z-scores standardised across samples and relative to mean expression (red = expression above the mean; blue expression below the mean). On the horizontal axis patients were arranged according to survival (red = death during ICU stay; orange = discharged alive from ICU but died during hospital stay; yellow = discharged alive from hospital but died within six months; blue = alive at end of study [six months]).

**Table 38 Values for survival according to patient cluster: individuals identified at high risk for early death vs. others.**

| High risk (cluster 1) | | | | | | |
|---|---|---|---|---|---|---|
| **Time** | **n.risk** | **n.event** | **Survival** | **Std.err** | **Lower 95% CI** | **Upper 95% CI** |
| 1 | 20 | 1 | 0.95 | 0.0487 | 0.859 | 1 |
| 3 | 19 | 2 | 0.85 | 0.0798 | 0.707 | 1 |
| 4 | 17 | 3 | 0.7 | 0.1025 | 0.525 | 0.933 |
| 6 | 14 | 1 | 0.65 | 0.1067 | 0.471 | 0.897 |
| 8 | 13 | 1 | 0.6 | 0.1095 | 0.42 | 0.858 |
| 33 | 12 | 1 | 0.55 | 0.1112 | 0.37 | 0.818 |
| 42 | 11 | 1 | 0.5 | 0.1118 | 0.323 | 0.775 |
| 50 | 10 | 1 | 0.45 | 0.1112 | 0.277 | 0.731 |
| 60 | 9 | 1 | 0.4 | 0.1095 | 0.234 | 0.684 |
| 132 | 8 | 1 | 0.35 | 0.1067 | 0.193 | 0.636 |
| **Other clusters** | | | | | | |
| 0 | 116 | 2 | 0.983 | 0.0121 | 0.959 | 1 |
| 1 | 114 | 1 | 0.974 | 0.0147 | 0.946 | 1 |
| 2 | 113 | 1 | 0.966 | 0.0169 | 0.933 | 0.999 |
| 3 | 112 | 1 | 0.957 | 0.0189 | 0.921 | 0.995 |
| 5 | 111 | 3 | 0.931 | 0.0235 | 0.886 | 0.978 |
| 6 | 108 | 1 | 0.922 | 0.0248 | 0.875 | 0.972 |
| 9 | 107 | 1 | 0.914 | 0.0261 | 0.864 | 0.966 |
| 10 | 106 | 1 | 0.905 | 0.0272 | 0.853 | 0.96 |
| 14 | 105 | 1 | 0.897 | 0.0283 | 0.843 | 0.954 |
| 15 | 104 | 1 | 0.888 | 0.0293 | 0.832 | 0.947 |
| 17 | 103 | 1 | 0.879 | 0.0302 | 0.822 | 0.941 |
| 19 | 102 | 2 | 0.862 | 0.032 | 0.802 | 0.927 |
| 24 | 100 | 1 | 0.853 | 0.0328 | 0.791 | 0.92 |
| 25 | 99 | 1 | 0.845 | 0.0336 | 0.781 | 0.913 |
| 27 | 98 | 1 | 0.836 | 0.0344 | 0.771 | 0.906 |

| | | | | | | |
|---|---|---|---|---|---|---|
| 37 | 97 | 1 | 0.828 | 0.0351 | 0.762 | 0.899 |
| 38 | 96 | 1 | 0.819 | 0.0358 | 0.752 | 0.892 |
| 45 | 95 | 1 | 0.81 | 0.0364 | 0.742 | 0.885 |
| 46 | 94 | 1 | 0.802 | 0.037 | 0.732 | 0.878 |
| 78 | 93 | 1 | 0.793 | 0.0376 | 0.723 | 0.87 |
| 84 | 92 | 1 | 0.784 | 0.0382 | 0.713 | 0.863 |
| 99 | 91 | 1 | 0.776 | 0.0387 | 0.704 | 0.856 |
| 109 | 90 | 1 | 0.767 | 0.0392 | 0.694 | 0.848 |
| 143 | 89 | 1 | 0.759 | 0.0397 | 0.685 | 0.841 |
| 143 | 82 | 1 | 0.757 | 0.0415 | 0.68 | 0.843 |

Values for Kaplan Meier plot, see Figure 28 in chapter five. Time is measured in days up to 180; n.risk = number of individuals at risk of death; n.event = number of individuals who died at that time-point; Std.err = standard error; CI = confidence interval.

**Table 39 Pathways of leukocyte expressed genes in severe sepsis and septic shock patients correlated to plasma 25(OH)D.**

| Canonical Pathways | B-H p-value | Ratio | Molecules |
|---|---|---|---|
| EIF2 Signaling | 1.3E-12 | 0.16 | RPL11,SOS2,EIF2A,RPS11,RPL10A,RPL6,RPS27,RPL18A,RPL7A,EIF3A,RPL35,RPL12,PIK3R2,RPL18,PIK3C2B,RPL4,EIF3H,RPL3,RPS2,RPS19,EIF3E,EIF3G,RPL5,RPS5,EIF2B5,RPS15,RPL13A,RPS14,EIF3K |
| Regulation of eIF4 and p70S6K Signaling | 1.3E-05 | 0.11 | PIK3C2B,EIF3H,RPS2,SOS2,RPS19,EIF3E,EIF2A,RPS11,EIF3G,RPS27,RPS5,MKNK1,RPS15,EIF2B5,EIF3A,PIK3R2,RPS14,EIF3K |
| mTOR Signaling | 3.9E-04 | 0.10 | PIK3C2B,PRKCQ,EIF3H,RPS2,RPS6KA3,RPS19,EIF3E,RPS11,EIF3G,RPS27,RHOG,RPS5,PROK1,EIF3A,RPS15,PIK3R2,RPS14,EIF3K |
| tRNA Charging | 6.9E-03 | 0.18 | NARS,LARS,CARS,FARS2,KARS,AARS2,EPRS |
| T Cell Receptor Signaling | 2.1E-02 | 0.10 | CD247,PIK3C2B,CD28,PRKCQ,PAG1,RASGRP1,SOS2,PIK3R2,CARD11,ITK |
| iCOS-iCOSL Signaling in T Helper Cells | 2.5E-02 | 0.09 | CD247,PIK3C2B,CD28,PRKCQ,ITPR3,FCER1G,PIK3R2,IL2RB,PTEN,ITK |
| NF-κB Signaling | 2.5E-02 | 0.08 | TLR2,TLR4,PIK3C2B,PRKCQ,FCER1G,IGF1R,TLR8,PIK3R2,IRAK3,CARD11,MAP3K3,IRAK4,IRAK1 |
| Toll-like Receptor Signaling | 3.7E-02 | 0.12 | TLR2,TLR4,TOLLIP,TLR8,IRAK3,IRAK4,IRAK1 |
| iNOS Signaling | 4.2E-02 | 0.13 | TLR4,IFNGR1,JAK2,IRAK3,IRAK4,IRAK1 |
| IL-12 Signaling and Production in Macrophages | 4.2E-02 | 0.08 | TLR2,TLR4,PIK3C2B,PRKCQ,ORM1,ORM2,IFNGR1,CEBPB,PIK3R2,REL,SPI1 |
| Role of JAK1 and JAK3 in γc Cytokine Signaling | 4.6E-02 | 0.11 | IL7R,PIK3C2B,IRS2,PIK3R2,JAK2,STAT5B,IL2RB |

| | | | |
|---|---|---|---|
| Production of Nitric Oxide and Reactive Oxygen Species in Macrophages | 4.7E-02 | 0.07 | PIK3C2B,PRKCQ,IFNGR1,JAK2,SPI1,TLR2, TLR4,PPP1R3D,RHOG,ORM1,ORM2,PIK3R2 ,MAP3K3 |

P-values are adjusted for multiple tests, only significant pathways are shown (adjusted p-value <0.05). B-H = Benjamini-Hochberg.

**Table 40 Pathways of leukocyte expressed genes in severe sepsis and septic shock patients correlated to plasma cathelicidin.**

| Canonical Pathways | B-H p-value | Ratio | Molecules |
|---|---|---|---|
| CD28 Signaling in T Helper Cells | 3.39E-04 | 0.15 | CD247,PRKCQ,CD3E,NFKBIE,CD4,CHP1,A RPC5,PPP3CC,CD3G,NFKBIA,PPP3CB,PIK3 C3,ITPR3,FCER1G,CHUK,CARD11,MAP2K1 ,ITK |
| PKCθ Signaling in T Lymphocytes | 3.39E-04 | 0.14 | CD247,PRKCQ,CD3E,NFKBIE,CD4,CHP1,SO S2,PPP3CC,MAP3K4,CD3G,NFKBIA,PPP3C B,PIK3C3,FCER1G,MAP3K8,CHUK,CARD11 ,MAP3K3 |
| iCOS-iCOSL Signaling in T Helper Cells | 3.39E-04 | 0.15 | CD247,PRKCQ,CD3E,NFKBIE,CD4,CHP1,PP P3CC,IL2RB,PTEN,CD3G,NFKBIA,PPP3CB, PIK3C3,ITPR3,FCER1G,CHUK,ITK |
| T Cell Receptor Signaling | 6.46E-04 | 0.16 | CD247,PRKCQ,CD3E,CD4,SOS2,PPP3CC,CD 3G,NFKBIA,PPP3CB,RASGRP1,PAG1,PIK3C 3,CHUK,CARD11,MAP2K1,ITK |
| Role of NFAT in Regulation of the Immune Response | 1.51E-03 | 0.11 | CD247,PRKCQ,CD3E,CD79B,NFKBIE,CD4,C HP1,SOS2,FCER1A,PPP3CC,GNG5,CD3G,G NAI3,NFKBIA,PPP3CB,PIK3C3,ITPR3,FCER 1G,CHUK,MAP2K1,ITK |
| Calcium-induced T Lymphocyte Apoptosis | 1.00E-02 | 0.16 | CD247,CD3G,PRKCQ,PPP3CB,CD3E,CD4,IT PR3,CHP1,FCER1G,PPP3CC |

| | | | |
|---|---|---|---|
| Regulation of IL-2 Expression in Activated and Anergic T Lymphocytes | 1.12E-02 | 0.14 | CD247,CD3G,NFKBIA,PPP3CB,CD3E,NFKBIE,SOS2,CHP1,PPP3CC,CHUK,MAP2K1,CARD11 |

P-values are adjusted for multiple tests, only significant pathways are shown (adjusted p-value <0.05). B-H = Benjamini-Hochberg.

**Table 41 Pathways of leukocyte expressed genes in severe sepsis and septic shock patients correlated to polynucleocyte count.**

| Canonical Pathways | B-H p-value | Ratio | Molecules |
|---|---|---|---|
| Spermine and Spermidine Degradation I | 0.47 | 0.50 | SAT1,SAT2 |
| Granzyme B Signaling | 0.47 | 0.19 | PRF1,NUMA1,GZMB |
| Putrescine Degradation III | 0.47 | 0.19 | SAT1,SAT2,MAOA |
| Natural Killer Cell Signaling | 0.70 | 0.06 | CD247,RAF1,PAK2,KLRB1,SH2D1B,CD300A,PRKCA |
| Cytotoxic T Lymphocyte-mediated Apoptosis of Target Cells | 0.75 | 0.08 | CD247,PRF1,FASLG,GZMB |

P-values are adjusted for multiple tests, no pathways were significant after p-value correction for multiple tests (top pathways shown). B-H = Benjamini-Hochberg.

# REFERENCES


Abecasis, G. R., Auton, A., Brooks, L. D., DePristo, M. A., Durbin, R. M., Handsaker, R. E., Kang, H. M., et al. (**2012**). "An integrated map of genetic variation from 1,092 human genomes." **Nature** 491(7422): 56-65.

Adam, N., Kandelman, S., Mantz, J., Chretien, F. and Sharshar, T. (**2013**). "Sepsis-induced brain dysfunction." **Expert Rev Anti Infect Ther** 11(2): 211-221.

Adorini, L. and Penna, G. (**2008**). "Control of autoimmune diseases by the vitamin D endocrine system." **Nat Clin Pract Rheumatol** 4(8): 404-412.

Adorini, L. and Penna, G. (**2009**). "Induction of tolerogenic dendritic cells by vitamin D receptor agonists." **Handb Exp Pharmacol**(188): 251-273.

Ahn, J., Yu, K., Stolzenberg-Solomon, R., Simon, K. C., McCullough, M. L., Gallicchio, L., Jacobs, E. J., et al. (**2010**). "Genome-wide association study of circulating vitamin D levels." **Hum Mol Genet** 19(13): 2739-2745.

Alekseyenko, A. A., Peng, S., Larschan, E., Gorchakov, A. A., Lee, O. K., Kharchenko, P., McGrath, S. D., et al. (**2008**). "A sequence motif within chromatin entry sites directs MSL establishment on the Drosophila X chromosome." **Cell** 134(4): 599-609.

Alloza, I., Otaegui, D., de Lapuente, A. L., Antiguedad, A., Varade, J., Nunez, C., Arroyo, R., et al. (**2012**). "ANKRD55 and DHCR7 are novel multiple sclerosis risk loci." **Genes Immun** 13(3): 253-257.

Alonso, A. and Hernan, M. A. (**2008**). "Temporal trends in the incidence of multiple sclerosis: a systematic review." **Neurology** 71(18606967): 129-135.

Alroy, I., Towers, T. L. and Freedman, L. P. (**1995**). "Transcriptional repression of the interleukin-2 gene by vitamin D3: direct inhibition of NFATp/AP-1 complex formation by a nuclear hormone receptor." **Mol Cell Biol** 15(7565732): 5789-5799.

Amrein, K., Sourij, H., Wagner, G., Holl, A., Pieber, T. R., Smolle, K. H., Stojakovic, T., et al. (**2011**). "Short-term effects of high-dose oral vitamin D3 in critically ill vitamin D deficient patients: a randomized, double-blind, placebo-controlled pilot study." **Crit Care** 15(2): R104.

Andraos, C., Koorsen, G., Knight, J. C. and Bornman, L. (**2011**). "Vitamin D receptor gene methylation is associated with ethnicity, tuberculosis, and TaqI polymorphism." **Hum Immunol** 72(21168462): 262-268.

Angus, D. C., Marrie, T. J., Obrosky, D. S., Clermont, G., Dremsizov, T. T., Coley, C., Fine, M. J., et al. (**2002**). "Severe community-acquired pneumonia: use of intensive care services and evaluation of American and British Thoracic Society Diagnostic criteria." **Am J Respir Crit Care Med** 166(5): 717-723.

Annane, D. (**2009**). "Improving clinical trials in the critically ill: unique challenge--sepsis." **Crit Care Med** 37(1 Suppl): S117-128.

Annane, D., Bellissant, E. and Cavaillon, J. M. (**2005**). "Septic shock." **Lancet** 365(9453): 63-78.

Antico, A., Tampoia, M., Tozzoli, R. and Bizzaro, N. (**2012**). "Can supplementation with vitamin D reduce the risk or modify the course of autoimmune diseases? A systematic review of the literature." **Autoimmun Rev** 12(2): 127-136.

ANZgene (**2009**). "Genome-wide association study identifies new multiple sclerosis susceptibility loci on chromosomes 12 and 20." **Nat Genet** 41(7): 824-828.

Aris, R. M., Merkel, P. A., Bachrach, L. K., Borowitz, D. S., Boyle, M. P., Elkin, S. L., Guise, T. A., et al. (**2005**). "Guide to bone health and disease in cystic fibrosis." **J Clin Endocrinol Metab** 90(15613415): 1888-1896.

Armas, L. A., Hollis, B. W. and Heaney, R. P. (**2004**). "Vitamin D2 is much less effective than vitamin D3 in humans." **J Clin Endocrinol Metab** 89(11): 5387-5391.

Arnon, R. and Aharoni, R. (**2004**). "Mechanism of action of glatiramer acetate in multiple sclerosis and its potential for the development of new applications." **Proc Natl Acad Sci U S A** 101 Suppl 2: 14593-14598.

Ascherio, A. and Munger, K. L. (**2007**). "Environmental risk factors for multiple sclerosis. Part I: the role of infection." **Ann Neurol** 61(4): 288-299.

Ascherio, A. and Munger, K. L. (**2007**). "Environmental risk factors for multiple sclerosis. Part II: Noninfectious factors." **Ann Neurol** 61(6): 504-513.

Ascherio, A., Munger, K. L. and Lunemann, J. D. (**2012**). "The initiation and prevention of multiple sclerosis." **Nat Rev Neurol** 8(11): 602-612.

Autier, P. and Gandini, S. (**2007**). "Vitamin D supplementation and total mortality: a meta-analysis of randomized controlled trials." **Arch Intern Med** 167(16): 1730-1737.

Badenhoop, K., Kahles, H. and Penna-Martinez, M. (**2012**). "Vitamin D, immune tolerance, and prevention of type 1 diabetes." **Curr Diab Rep** 12(6): 635-642.

Baeke, F., Korf, H., Overbergh, L., van Etten, E., Verstuyf, A., Gysemans, C. and Mathieu, C. (**2010**). "Human T lymphocytes are direct targets of 1,25-dihydroxyvitamin D3 in the immune system." **J Steroid Biochem Mol Biol** 121(1-2): 221-227.

Baeke, F., Takiishi, T., Korf, H., Gysemans, C. and Mathieu, C. (**2010**). "Vitamin D: modulator of the immune system." **Curr Opin Pharmacol** 10(4): 482-496.

Banchereau, J., Pascual, V. and O'Garra, A. (**2012**). "From IL-2 to IL-37: the expanding spectrum of anti-inflammatory cytokines." **Nat Immunol** 13(10): 925-931.

Bang, U. C., Novovic, S., Andersen, A. M., Fenger, M., Hansen, M. B. and Jensen, J. E. (**2011**). "Variations in serum 25-hydroxyvitamin D during acute pancreatitis: an exploratory longitudinal study." **Endocr Res** 36(4): 135-141.

Barnay-Verdier, S., Fattoum, L., Borde, C., Kaveri, S., Gibot, S. and Marechal, V. (**2011**). "Emergence of autoantibodies to HMGB1 is associated with survival in patients with septic shock." **Intensive Care Med** 37(6): 957-962.

Beecham, A. H., Patsopoulos, N. A., Xifara, D. K., Davis, M. F., Kemppinen, A., Cotsapas, C., Shah, T. S., et al. (**2013**). "Analysis of immune-related loci identifies 48 new susceptibility variants for multiple sclerosis." **Nat Genet**.

Berlanga-Taylor, A. J., Disanto, G., Ebers, G. C. and Ramagopalan, S. V. (**2011**). "Vitamin D-gene interactions in multiple sclerosis." **J Neurol Sci** 311(1-2): 32-36.

Berlanga-Taylor, A. J. and Ramagopalan, S. V. (**2013**). "Vitamin D and multiple sclerosis: what is the clinical impact?" **Expert Opin Med Diagn**.

Bertoldo, F., Pancheri, S., Zenari, S., Boldini, S., Giovanazzi, B., Zanatta, M., Valenti, M. T., et al. (**2010**). "Serum 25-hydroxyvitamin D levels modulate the acute-phase response associated with the first nitrogen-containing bisphosphonate infusion." **J Bone Miner Res** 25(3): 447-454.

Bischoff-Ferrari, H. A., Giovannucci, E., Willett, W. C., Dietrich, T. and Dawson-Hughes, B. (**2006**). "Estimation of optimal serum concentrations of 25-hydroxyvitamin D for multiple health outcomes." **Am J Clin Nutr** 84(1): 18-28.

Bone, R. C. (**1996**). "Immunologic dissonance: a continuing evolution in our understanding of the systemic inflammatory response syndrome (SIRS) and the multiple organ dysfunction syndrome (MODS)." **Ann Intern Med** 125(8): 680-687.

Boomer, J. S., To, K., Chang, K. C., Takasu, O., Osborne, D. F., Walton, A. H., Bricker, T. L., et al. (**2011**). "Immunosuppression in patients who die of sepsis and multiple organ failure." **JAMA** 306(23): 2594-2605.

Bosmann, M. and Ward, P. A. (**2013**). "The inflammatory response in sepsis." **Trends Immunol** 34(3): 129-136.

Bouchon, A., Facchetti, F., Weigand, M. A. and Colonna, M. (**2001**). "TREM-1 amplifies inflammation and is a crucial mediator of septic shock." **Nature** 410(6832): 1103-1107.

Bouillon, R., Bischoff-Ferrari, H. and Willett, W. (**2008**). "Vitamin D and health: perspectives from mice and man." **J Bone Miner Res** 23(7): 974-979.

Bouillon, R., Carmeliet, G., Verlinden, L., van Etten, E., Verstuyf, A., Luderer, H. F., Lieben, L., et al. (**2008**). "Vitamin D and human health: lessons from vitamin D receptor null mice." **Endocr Rev** 29(6): 726-776.

Boyle, A. P., Hong, E. L., Hariharan, M., Cheng, Y., Schaub, M. A., Kasowski, M., Karczewski, K. J., et al. (**2012**). "Annotation of functional variation in personal genomes using RegulomeDB." **Genome Res** 22(9): 1790-1797.

Braun, A., Chang, D., Mahadevappa, K., Gibbons, F. K., Liu, Y., Giovannucci, E. and Christopher, K. B. (**2011**). "Association of low serum 25-hydroxyvitamin D levels and mortality in the critically ill." **Crit Care Med** 39(4): 671-677.

Braun, A. B., Gibbons, F. K., Litonjua, A. A., Giovannucci, E. and Christopher, K. B. (**2011**). "Low serum 25-hydroxyvitamin D at critical care initiation is associated with increased mortality." **Crit Care Med**.

Brealey, D., Brand, M., Hargreaves, I., Heales, S., Land, J., Smolenski, R., Davies, N. A., et al. (**2002**). "Association between mitochondrial dysfunction and severity and outcome of septic shock." **Lancet** 360(9328): 219-223.

Brien, G. L., Gambero, G., O'Connell, D. J., Jerman, E., Turner, S. A., Egan, C. M., Dunne, E. J., et al. (**2012**). "Polycomb PHF19 binds H3K36me3 and recruits PRC2 and demethylase NO66 to embryonic stem cell genes during differentiation." **Nat Struct Mol Biol** 19(12): 1273-1281.

Bronnum-Hansen, H., Koch-Henriksen, N. and Stenager, E. (**2004**). "Trends in survival and cause of death in Danish patients with multiple sclerosis." **Brain** 127(Pt 4): 844-850.

Brown, K. A., Brain, S. D., Pearson, J. D., Edgeworth, J. D., Lewis, S. M. and Treacher, D. F. (**2006**). "Neutrophils in development of multiple organ failure in sepsis." **Lancet** 368(9530): 157-169.

Brumbaugh, P. F. and Haussler, M. R. (**1975**). "Specific binding of 1alpha,25-dihydroxycholecalciferol to nuclear components of chick intestine." **J Biol Chem** 250(163254): 1588-1594.

Brun-Buisson, C., Doyon, F. and Carlet, J. (**1996**). "Bacteremia and severe sepsis in adults: a multicenter prospective survey in ICUs and wards of 24 hospitals. French Bacteremia-Sepsis Study Group." **Am J Respir Crit Care Med** 154(3 Pt 1): 617-624.

Buratowski, S. and Chodosh, L. A. (**2001**). "Mobility shift DNA-binding assay using gel electrophoresis." **Curr Protoc Mol Biol** Chapter 12: Unit 12 12.

Burton, P. R., Clayton, D. G., Cardon, L. R., Craddock, N., Deloukas, P., Duncanson, A., Kwiatkowski, D. P., et al. (**2007**). "Association scan of 14,500 nonsynonymous SNPs in four diseases identifies autoimmunity variants." **Nat Genet** 39(11): 1329-1337.

Carlberg, C. and Campbell, M. J. (**2013**). "Vitamin D receptor signaling mechanisms: integrated actions of a well-defined transcription factor." **Steroids** 78(2): 127-136.

Carlson, N. G. and Rose, J. W. (**2013**). "Vitamin D as a clinical biomarker in multiple sclerosis." **Expert Opinion on Medical Diagnostics**.

Cecchi, A., Bonizzoli, M., Douar, S., Mangini, M., Paladini, S., Gazzini, B., Degl'innocenti, S., et al. (**2011**). "Vitamin D deficiency in septic patients at ICU admission is not a mortality predictor." **Minerva Anestesiol** 77(12): 1184-1189.

Chabner, B. A. (**2011**). "Early accelerated approval for highly targeted cancer drugs." **N Engl J Med** 364(12): 1087-1089.

Chen, H., Hewison, M., Hu, B. and Adams, J. S. (**2003**). "Heterogeneous nuclear ribonucleoprotein (hnRNP) binding to hormone response elements: a cause of vitamin D resistance." **Proc Natl Acad Sci U S A** 100(12716975): 6109-6114.

Chen, S., Sims, G. P., Chen, X. X., Gu, Y. Y., Chen, S. and Lipsky, P. E. (**2007**). "Modulatory effects of 1,25-dihydroxyvitamin D3 on human B cell differentiation." **J Immunol** 179(17641030): 1634-1647.

Cohen, J. (**2002**). "The immunopathogenesis of sepsis." **Nature** 420(6917): 885-891.

Cohen, J., Opal, S. and Calandra, T. (**2012**). "Sepsis studies need new direction." **Lancet Infect Dis** 12(7): 503-505.

Compston, A. and Coles, A. (**2008**). "Multiple sclerosis." **Lancet** 372(9648): 1502-1517.

Consortium, I. M. S. G. (**2009**). "The expanding genetic overlap between multiple sclerosis and type I diabetes." **Genes and Immunity** 10(1): 11-14.

Cook, D. N., Pisetsky, D. S. and Schwartz, D. A. (**2004**). "Toll-like receptors in the pathogenesis of human disease." **Nat Immunol** 5(10): 975-979.

Cookson, W., Liang, L., Abecasis, G., Moffatt, M. and Lathrop, M. (**2009**). "Mapping complex disease traits with global gene expression." **Nat Rev Genet** 10(3): 184-194.

Coopersmith, C. M., Stromberg, P. E., Dunne, W. M., Davis, C. G., Amiot, D. M., 2nd, Buchman, T. G., Karl, I. E., et al. (**2002**). "Inhibition of intestinal epithelial apoptosis and survival in a murine model of pneumonia-induced sepsis." **JAMA** 287(13): 1716-1721.

D'Ambrosio, D., Cippitelli, M., Cocciolo, M. G., Mazzeo, D., Di Lucia, P., Lang, R., Sinigaglia, F., et al. (**1998**). "Inhibition of IL-12 production by 1,25-dihydroxyvitamin D3. Involvement of NF-kappaB downregulation in transcriptional repression of the p40 gene." **J Clin Invest** 101(1): 252-262.

D'Netto, M. J., Ward, H., Morrison, K. M., Ramagopalan, S. V., Dyment, D. A., DeLuca, G. C., Handunnetthi, L., et al. (**2009**). "Risk alleles for multiple sclerosis in multiplex families." **Neurology** 72(23): 1984-1988.

Danai, P. A., Sinha, S., Moss, M., Haber, M. J. and Martin, G. S. (**2007**). "Seasonal variation in the epidemiology of sepsis." **Crit Care Med** 35(2): 410-415.

De Jager, P. L., Jia, X., Wang, J., de Bakker, P. I., Ottoboni, L., Aggarwal, N. T., Piccio, L., et al. (**2009**). "Meta-analysis of genome scans and replication identify CD6, IRF8 and TNFRSF1A as new multiple sclerosis susceptibility loci." **Nat Genet** 41(7): 776-782.

De Jonghe, B., Sharshar, T., Lefaucheur, J. P., Authier, F. J., Durand-Zaleski, I., Boussarsar, M., Cerf, C., et al. (**2002**). "Paresis acquired in the intensive care unit: a prospective multicenter study." **JAMA** 288(22): 2859-2867.

Decard, B. F., von Ahsen, N., Grunwald, T., Streit, F., Stroet, A., Niggemeier, P., Schottstedt, V., et al. (**2012**). "Low vitamin D and elevated immunoreactivity against Epstein-Barr virus before first clinical manifestation of multiple sclerosis." **J Neurol Neurosurg Psychiatry** 83(12): 1170-1173.

Degenhardt, A., Ramagopalan, S. V., Scalfari, A. and Ebers, G. C. (**2009**). "Clinical prognostic factors in multiple sclerosis: a natural history review." **Nat Rev Neurol** 5(12): 672-682.

Dellinger, R. P., Levy, M. M., Rhodes, A., Annane, D., Gerlach, H., Opal, S. M., Sevransky, J. E., et al. (**2013**). "Surviving sepsis campaign: international guidelines for management of severe sepsis and septic shock: 2012." **Crit Care Med** 41(2): 580-637.

DeLuca, H. F. (**2004**). "Overview of general physiologic features and functions of vitamin D." **Am J Clin Nutr** 80(6 Suppl): 1689S-1696S.

Ding, N., Yu, R. T., Subramaniam, N., Sherman, M. H., Wilson, C., Rao, R., Leblanc, M., et al. (**2013**). "A vitamin D receptor/SMAD genomic circuit gates hepatic fibrotic response." **Cell** 153(3): 601-613.

Disanto, G., Hall, C., Lucas, R., Ponsonby, A. L., Berlanga-Taylor, A. J., Giovannoni, G. and Ramagopalan, S. V. (**2013**). "Assessing interactions between HLA-DRB1*15 and infectious mononucleosis on the risk of multiple sclerosis." **Mult Scler**.

Doig, C. J., Sutherland, L. R., Sandham, J. D., Fick, G. H., Verhoef, M. and Meddings, J. B. (**1998**). "Increased intestinal permeability is associated with the development of multiple organ dysfunction syndrome in critically ill ICU patients." **Am J Respir Crit Care Med** 158(2): 444-451.

Du, P., Kibbe, W. A. and Lin, S. M. (**2008**). "lumi: a pipeline for processing Illumina microarray." **Bioinformatics** 24(13): 1547-1548.

Dunham, I., Kundaje, A., Aldred, S. F., Collins, P. J., Davis, C. A., Doyle, F., Epstein, C. B., et al. (**2012**). "An integrated encyclopedia of DNA elements in the human genome." **Nature** 489(7414): 57-74.

Dyment, D. A., Herrera, B. M., Cader, M. Z., Willer, C. J., Lincoln, M. R., Sadovnick, A. D., Risch, N., et al. (**2005**). "Complex interactions among MHC haplotypes in multiple sclerosis: susceptibility and resistance." **Hum Mol Genet** 14(14): 2019-2026.

Ebers, G. (**2010**). "Month of birth and multiple sclerosis risk in Scotland." **Eur Neurol** 63(1): 41-42.

Ebers, G. C. (**2008**). "Environmental factors and multiple sclerosis." **Lancet Neurol** 7(3): 268-277.

Ebers, G. C., Kukay, K., Bulman, D. E., Sadovnick, A. D., Rice, G., Anderson, C., Armstrong, H., et al. (**1996**). "A full genome search in multiple sclerosis." **Nat Genet** 13(4): 472-476.

Edfeldt, K., Liu, P. T., Chun, R., Fabri, M., Schenk, M., Wheelwright, M., Keegan, C., et al. (**2010**). "T-cell cytokines differentially control human monocyte antimicrobial responses by regulating vitamin D metabolism." **Proc Natl Acad Sci U S A** 107(52): 22593-22598.

Etzioni, A., Hochberg, Z., Pollak, S., Meshulam, T., Zakut, V., Tzehoval, E., Keisari, Y., et al. (**1989**). "Defective leukocyte fungicidal activity in end-organ resistance to 1,25-dihydroxyvitamin D." **Pediatr Res** 25(3): 276-279.

Eyles, D. W., Smith, S., Kinobe, R., Hewison, M. and McGrath, J. J. (**2005**). "Distribution of the vitamin D receptor and 1 alpha-hydroxylase in human brain." **J Chem Neuroanat** 29(15589699): 21-30.

Fabri, M., Stenger, S., Shin, D. M., Yuk, J. M., Liu, P. T., Realegeno, S., Lee, H. M., et al. (**2011**). "Vitamin D is required for IFN-gamma-mediated antimicrobial activity of human macrophages." **Sci Transl Med** 3(104): 104ra102.

Fairfax, B. P., Makino, S., Radhakrishnan, J., Plant, K., Leslie, S., Dilthey, A., Ellis, P., et al. (**2012**). "Genetics of gene expression in primary immune cells identifies cell type-specific master regulators and roles of HLA alleles." **Nat Genet**.

Fillatreau, S., Sweenie, C. H., McGeachy, M. J., Gray, D. and Anderton, S. M. (**2002**). "B cells regulate autoimmunity by provision of IL-10." **Nat Immunol** 3(12244307): 944-950.

Fraser, D. R. and Kodicek, E. (**1970**). "Unique biosynthesis by kidney of a biological active vitamin D metabolite." **Nature** 228(4319631): 764-766.

Frohman, E. M., Racke, M. K. and Raine, C. S. (**2006**). "Multiple sclerosis--the plaque and its pathogenesis." **N Engl J Med** 354(9): 942-955.

Fujiki, R., Kim, M. S., Sasaki, Y., Yoshimura, K., Kitagawa, H. and Kato, S. (**2005**). "Ligand-induced transrepression by VDR through association of WSTF with acetylated histones." **EMBO J** 24(22): 3881-3894.

Fullwood, M. J., Liu, M. H., Pan, Y. F., Liu, J., Xu, H., Mohamed, Y. B., Orlov, Y. L., et al. (**2009**). "An oestrogen-receptor-alpha-bound human chromatin interactome." **Nature** 462(7269): 58-64.

Gandhi, R., Laroni, A. and Weiner, H. L. (**2010**). "Role of the innate immune system in the pathogenesis of multiple sclerosis." **J Neuroimmunol** 221(1-2): 7-14.

Ghazalpour, A., Bennett, B., Petyuk, V. A., Orozco, L., Hagopian, R., Mungrue, I. N., Farber, C. R., et al. (**2011**). "Comparative analysis of proteome and transcriptome variation in mouse." **PLoS Genet** 7(6): e1001393.

Giacomini, E., Severa, M., Rizzo, F., Mechelli, R., Annibali, V., Ristori, G., Riccieri, V., et al. (**2013**). "IFN-beta therapy modulates B-cell and monocyte crosstalk via TLR7 in multiple sclerosis patients." **Eur J Immunol**.

Giamarellos-Bourboulis, E. J. (**2013**). "The failure of biologics in sepsis: where do we stand?" **Int J Antimicrob Agents** 42 Suppl: S45-47.

Ginde, A. A., Camargo, C. A., Jr. and Shapiro, N. I. (**2011**). "Vitamin D insufficiency and sepsis severity in emergency department patients with suspected infection." **Acad Emerg Med** 18(5): 551-554.

Giovannoni, G. and Ebers, G. (**2007**). "Multiple sclerosis: the environment and causation." **Curr Opin Neurol** 20(3): 261-268.

Gold, R., Giovannoni, G., Selmaj, K., Havrdova, E., Montalban, X., Radue, E. W., Stefoski, D., et al. (**2013**). "Daclizumab high-yield process in relapsing-remitting multiple sclerosis (SELECT): a randomised, double-blind, placebo-controlled trial." **Lancet**.

Gombart, A. F. (**2009**). "The vitamin D-antimicrobial peptide pathway and its role in protection against infection." **Future Microbiol** 4(19895218): 1151-1165.

Goodin, D. S., Reder, A. T., Ebers, G. C., Cutter, G., Kremenchutzky, M., Oger, J., Langdon, D., et al. (**2012**). "Survival in MS: a randomized cohort study 21 years after the start of the pivotal IFNbeta-1b trial." **Neurology** 78(17): 1315-1322.

Gorman, S., Judge, M. A. and Hart, P. H. (**2010**). "Gene regulation by 1,25-dihydroxyvitamin D3 in CD4+CD25+ cells is enabled by IL-2." **J Invest Dermatol** 130(10): 2368-2376.

Green, N. M. and Marshak-Rothstein, A. (**2011**). "Toll-like receptor driven B cell activation in the induction of systemic autoimmunity." **Semin Immunol** 23(2): 106-112.

Gregori, S., Casorati, M., Amuchastegui, S., Smiroldo, S., Davalli, A. M. and Adorini, L. (**2001**). "Regulatory T cells induced by 1 alpha,25-dihydroxyvitamin D3 and mycophenolate mofetil treatment mediate transplantation tolerance." **J Immunol** 167(11489974): 1945-1953.

Grossman, S. R., Shlyakhter, I., Karlsson, E. K., Byrne, E. H., Morales, S., Frieden, G., Hostetter, E., et al. (**2010**). "A composite of multiple signals distinguishes causal variants in regions of positive selection." **Science** 327(5967): 883-886.

Grossniklaus, U., Kelly, B., Ferguson-Smith, A. C., Pembrey, M. and Lindquist, S. (**2013**). "Transgenerational epigenetic inheritance: how important is it?" **Nat Rev Genet** 14(3): 228-235.

Guntupalli, K., Dean, N., Morris, P. E., Bandi, V., Margolis, B., Rivers, E., Levy, M., et al. (**2013**). "A phase 2 randomized, double-blind, placebo-controlled study of the safety and efficacy of talactoferrin in patients with severe sepsis." **Crit Care Med** 41(3): 706-716.

Gwinn, M., Guessous, I. and Khoury, M. J. (**2009**). "Invited commentary: genes, environment, and hybrid vigor." **Am J Epidemiol** 170(6): 703-707.

Handel, A. E., Giovannoni, G., Ebers, G. C. and Ramagopalan, S. V. (**2010**). "Environmental factors and their timing in adult-onset multiple sclerosis." **Nat Rev Neurol** 6(3): 156-166.

Handel, A. E., Handunnetthi, L., Giovannoni, G., Ebers, G. C. and Ramagopalan, S. V. (**2010**). "Genetic and environmental factors and the distribution of multiple sclerosis in Europe." **Eur J Neurol** 17(9): 1210-1214.

Handunnetthi, L., Ramagopalan, S. V., Ebers, G. C. and Knight, J. C. (**2010**). "Regulation of major histocompatibility complex class II gene expression, genetic variation and disease." **Genes Immun** 11(2): 99-112.

Harrison, D. A., Welch, C. A. and Eddleston, J. M. (**2006**). "The epidemiology of severe sepsis in England, Wales and Northern Ireland, 1996 to 2004: secondary analysis of a high quality clinical database, the ICNARC Case Mix Programme Database." **Crit Care** 10(2): R42.

Harrois, A., Huet, O. and Duranteau, J. (**2009**). "Alterations of mitochondrial function in sepsis and critical illness." **Curr Opin Anaesthesiol** 22(2): 143-149.

Haussler, M. R., Jurutka, P. W., Mizwicki, M. and Norman, A. W. (**2011**). "Vitamin D receptor (VDR)-mediated actions of 1alpha,25(OH)(2)vitamin D(3): genomic and non-genomic mechanisms." **Best Pract Res Clin Endocrinol Metab** 25(4): 543-559.

Heger, A., Webber, C., Goodson, M., Ponting, C. P. and Lunter, G. (**2013**). "GAT: a simulation framework for testing the association of genomic intervals." **Bioinformatics**.

Heikkinen, S., Vaisanen, S., Pehkonen, P., Seuter, S., Benes, V. and Carlberg, C. (**2011**). "Nuclear hormone 1{alpha},25-dihydroxyvitamin D3 elicits a genome-wide shift in the locations of VDR chromatin occupancy." **Nucleic Acids Res**.

Heine, G., Niesner, U., Chang, H. D., Steinmeyer, A., Zugel, U., Zuberbier, T., Radbruch, A., et al. (**2008**). "1,25-dihydroxyvitamin D(3) promotes IL-10 production in human B cells." **Eur J Immunol** 38(8): 2210-2218.

Hernan, M. A., Jick, S. S., Logroscino, G., Olek, M. J., Ascherio, A. and Jick, H. (**2005**). "Cigarette smoking and the progression of multiple sclerosis." **Brain** 128(Pt 6): 1461-1465.

Hernan, M. A., Olek, M. J. and Ascherio, A. (**1999**). "Geographic variation of MS incidence in two prospective studies of US women." **Neurology** 53(10563617): 1711-1718.

Hewison, M. (**2011**). "Antibacterial effects of vitamin D." **Nat Rev Endocrinol** 7(6): 337-345.

Hill, A. B. (**1965**). "THE ENVIRONMENT AND DISEASE: ASSOCIATION OR CAUSATION?" **Proc R Soc Med** 58: 295-300.

Hindorff, L. A., Sethupathy, P., Junkins, H. A., Ramos, E. M., Mehta, J. P., Collins, F. S. and Manolio, T. A. (**2009**). "Potential etiologic and functional implications of genome-wide association loci for human diseases and traits." **Proc Natl Acad Sci U S A** 106(23): 9362-9367.

Holick, M. F. (**2007**). "Vitamin D deficiency." **N Engl J Med** 357(3): 266-281.

Holick, M. F. (**2011**). "Vitamin D: Evolutionary, Physiological and Health Perspectives." **Curr Drug Targets**.

Horiuchi, H., Nagata, I. and Komoriya, K. (**1991**). "Protective effect of vitamin D3 analogues on endotoxin shock in mice." **Agents Actions** 33(3-4): 343-348.

Horst, R. L., Reinhardt, T. A. and Reddy, G. S. (2005). Vitamin D metabolism. Boston, Elsevier Academic Press.

Hotchkiss, R. S., Monneret, G. and Payen, D. (**2013**). "Immunosuppression in sepsis: a novel understanding of the disorder and a new therapeutic approach." **Lancet Infect Dis** 13(3): 260-268.

Hsu, F., Kent, W. J., Clawson, H., Kuhn, R. M., Diekhans, M. and Haussler, D. (**2006**). "The UCSC Known Genes." **Bioinformatics** 22(16500937): 1036-1046.

Hudson, T. J., Anderson, W., Artez, A., Barker, A. D., Bell, C., Bernabe, R. R., Bhan, M. K., et al. (**2010**). "International network of cancer genome projects." **Nature** 464(7291): 993-998.

Hypponen, E. and Power, C. (**2007**). "Hypovitaminosis D in British adults at age 45 y: nationwide cohort study of dietary and lifestyle predictors." **Am J Clin Nutr** 85(3): 860-868.

Idaghdour, Y., Storey, J. D., Jadallah, S. J. and Gibson, G. (**2008**). "A genome-wide gene expression signature of environmental geography in leukocytes of Moroccan Amazighs." **PLoS Genet** 4(4): e1000052.

Islam, T., Gauderman, W. J., Cozen, W., Hamilton, A. S., Burnett, M. E. and Mack, T. M. (**2006**). "Differential twin concordance for multiple sclerosis by latitude of birthplace." **Ann Neurol** 60(1): 56-64.

Jablonski, N. G. and Chaplin, G. (**2000**). "The evolution of human skin coloration." **J Hum Evol** 39(1): 57-106.

Jeng, L., Yamshchikov, A. V., Judd, S. E., Blumberg, H. M., Martin, G. S., Ziegler, T. R. and Tangpricha, V. (**2009**). "Alterations in vitamin D status and anti-microbial peptide levels in patients in the intensive care unit with sepsis." **J Transl Med** 7: 28.

Ji, H., Li, X., Wang, Q. F. and Ning, Y. (**2013**). "Differential principal component analysis of ChIP-seq." **Proc Natl Acad Sci U S A** 110(17): 6789-6794.

Johnson, M. T., Reichley, R., Hoppe-Bauer, J., Dunne, W. M., Micek, S. and Kollef, M. (**2011**). "Impact of previous antibiotic therapy on outcome of Gram-negative severe sepsis." **Crit Care Med** 39(8): 1859-1865.

Jones, G. (**2008**). "Pharmacokinetics of vitamin D toxicity." **Am J Clin Nutr** 88(2): 582S-586S.

Kahn, H. S., Morgan, T. M., Case, L. D., Dabelea, D., Mayer-Davis, E. J., Lawrence, J. M., Marcovina, S. M., et al. (**2009**). "Association of type 1 diabetes with month of birth among U.S. youth: The SEARCH for Diabetes in Youth Study." **Diabetes Care** 32(11): 2010-2015.

Kakizawa, T., Miyamoto, T., Ichikawa, K., Kaneko, A., Suzuki, S., Hara, M., Nagasawa, T., et al. (**1999**). "Functional interaction between Oct-1 and retinoid X receptor." **J Biol Chem** 274(27): 19103-19108.

Kampman, M. T. and Brustad, M. (**2008**). "Vitamin D: a candidate for the environmental effect in multiple sclerosis - observations from Norway." **Neuroepidemiology** 30(3): 140-146.

Kampman, M. T., Wilsgaard, T. and Mellgren, S. I. (**2007**). "Outdoor activities and diet in childhood and adolescence relate to MS risk above the Arctic Circle." **J Neurol** 254(4): 471-477.

Kandoth, C., Schultz, N., Cherniack, A. D., Akbani, R., Liu, Y., Shen, H., Robertson, A. G., et al. (**2013**). "Integrated genomic characterization of endometrial carcinoma." **Nature** 497(7447): 67-73.

Karohl, C., Su, S., Kumari, M., Tangpricha, V., Veledar, E., Vaccarino, V. and Raggi, P. (**2010**). "Heritability and seasonal variability of vitamin D concentrations in male twins." **Am J Clin Nutr**.

Khoo, A. L., Chai, L. Y., Koenen, H. J., Kullberg, B. J., Joosten, I., van der Ven, A. J. and Netea, M. G. (**2011**). "1,25-dihydroxyvitamin D3 modulates cytokine production induced by Candida albicans: impact of seasonal variation of immune responses." **J Infect Dis** 203(1): 122-130.

Khoo, A. L., Chai, L. Y., Koenen, H. J., Oosting, M., Steinmeyer, A., Zuegel, U., Joosten, I., et al. (**2011**). "Vitamin D(3) down-regulates proinflammatory cytokine response to Mycobacterium tuberculosis through pattern recognition receptors while inducing protective cathelicidin production." **Cytokine** 55(2): 294-300.

Khoo, A. L., Chai, L. Y., Koenen, H. J., Sweep, F. C., Joosten, I., Netea, M. G. and van der Ven, A. J. (**2011**). "Regulation of cytokine responses by seasonality of vitamin D status in healthy individuals." **Clin Exp Immunol**.

Kim, S., Shevde, N. K. and Pike, J. W. (**2005**). "1,25-Dihydroxyvitamin D3 stimulates cyclic vitamin D receptor/retinoid X receptor DNA-binding, co-activator recruitment, and histone acetylation in intact osteoblasts." **J Bone Miner Res** 20(15647825): 305-317.

Kingwell, E., Marriott, J. J., Jette, N., Pringsheim, T., Makhani, N., Morrow, S. A., Fisk, J. D., et al. (**2013**). "Incidence and prevalence of multiple sclerosis in Europe: a systematic review." **BMC Neurol** 13(1): 128.

Kirschenbaum, L. A., Aziz, M., Astiz, M. E., Saha, D. C. and Rackow, E. C. (**2000**). "Influence of rheologic changes and platelet-neutrophil interactions on cell filtration in sepsis." **Am J Respir Crit Care Med** 161(5): 1602-1607.

Kitagawa, H., Fujiki, R., Yoshimura, K., Mezaki, Y., Uematsu, Y., Matsui, D., Ogawa, S., et al. (**2003**). "The chromatin-remodeling complex WINAC targets a nuclear receptor to promoters and is impaired in Williams syndrome." **Cell** 113(12837248): 905-917.

Kitanaka, S. (**2010**). "[Genetic basis for skeletal disease. Hereditary rickets]." **Clin Calcium** 20(8): 1238-1244.

Kleinewietfeld, M., Manzel, A., Titze, J., Kvakan, H., Yosef, N., Linker, R. A., Muller, D. N., et al. (**2013**). "Sodium chloride drives autoimmune disease by the induction of pathogenic TH17 cells." **Nature** 496(7446): 518-522.

Knight, J. C., Keating, B. J. and Kwiatkowski, D. P. (**2004**). "Allele-specific repression of lymphotoxin-alpha by activated B cell factor-1." **Nat Genet** 36(4): 394-399.

Knight, J. C., Keating, B. J., Rockett, K. A. and Kwiatkowski, D. P. (**2003**). "In vivo characterization of regulatory polymorphisms by allele-specific quantification of RNA polymerase loading." **Nat Genet** 33(4): 469-475.

Knight, J. C., Udalova, I., Hill, A. V., Greenwood, B. M., Peshu, N., Marsh, K. and Kwiatkowski, D. (**1999**). "A polymorphism that affects OCT-1 binding to the TNF promoter region is associated with severe malaria." **Nat Genet** 22(2): 145-150.

Koch-Henriksen, N., Stenager, E. and Bronnum-Hansen, H. (**2011**). "Studies based on the Danish Multiple Sclerosis Registry." **Scand J Public Health** 39(7 Suppl): 180-184.

Krishnan, A., Ochola, J., Mundy, J., Jones, M., Kruger, P., Duncan, E. and Venkatesh, B. (**2010**). "Acute fluid shifts influence the assessment of serum vitamin D status in critically ill patients." **Crit Care** 14(6): R216.

Kummer, S., von Gersdorff, G., Kemper, M. J. and Oh, J. (**2012**). "The influence of gender and sexual hormones on incidence and outcome of chronic kidney disease." **Pediatr Nephrol** 27(8): 1213-1219.

Kurtzke, J. F., Dean, G. and Botha, D. P. (**1970**). "A method for estimating the age at immigration of white immigrants to South Africa, with an example of its importance." **S Afr Med J** 44(23): 663-669.

Kurtzke, J. F., Delasnerie-Laupretre, N. and Wallin, M. T. (**1998**). "Multiple sclerosis in North African migrants to France." **Acta Neurol Scand** 98(5): 302-309.

Lange, N., Litonjua, A. A., Gibbons, F. K., Giovannucci, E. and Christopher, K. B. (**2013**). "Pre-hospital Vitamin D Concentration, Mortality, and Bloodstream Infection in a Hospitalized Patient Population." **Am J Med** 126(7): 640 e619-627.

Lappalainen, T., Sammeth, M., Friedlander, M. R., t Hoen, P. A., Monlong, J., Rivas, M. A., Gonzalez-Porta, M., et al. (**2013**). "Transcriptome and genome sequencing uncovers functional variation in humans." **Nature** 501(7468): 506-511.

Lee, P., Eisman, J. A. and Center, J. R. (**2009**). "Vitamin D deficiency in critically ill patients." **N Engl J Med** 360(18): 1912-1914.

Lemire, J. M. (**1995**). "Immunomodulatory actions of 1,25-dihydroxyvitamin D3." **J Steroid Biochem Mol Biol** 53(1-6): 599-602.

Lemire, J. M., Adams, J. S., Sakai, R. and Jordan, S. C. (**1984**). "1 alpha,25-dihydroxyvitamin D3 suppresses proliferation and immunoglobulin production by normal human peripheral blood mononuclear cells." **J Clin Invest** 74(2): 657-661.

Lescot, T., Karvellas, C., Beaussier, M. and Magder, S. (**2012**). "Acquired liver injury in the intensive care unit." **Anesthesiology** 117(4): 898-904.

Levy, M. M., Dellinger, R. P., Townsend, S. R., Linde-Zwirble, W. T., Marshall, J. C., Bion, J., Schorr, C., et al. (**2010**). "The Surviving Sepsis Campaign: results of an international guideline-based performance improvement program targeting severe sepsis." **Crit Care Med** 38(2): 367-374.

Levy, M. M., Fink, M. P., Marshall, J. C., Abraham, E., Angus, D., Cook, D., Cohen, J., et al. (**2003**). "2001 SCCM/ESICM/ACCP/ATS/SIS International Sepsis Definitions Conference." **Crit Care Med** 31(4): 1250-1256.

Lincoln, M. R., Montpetit, A., Cader, M. Z., Saarela, J., Dyment, D. A., Tiislar, M., Ferretti, V., et al. (**2005**). "A predominant role for the HLA class II region

in the association of the MHC region with multiple sclerosis." **Nat Genet** 37(10): 1108-1112.

Linker-Israeli, M., Elstner, E., Klinenberg, J. R., Wallace, D. J. and Koeffler, H. P. (**2001**). "Vitamin D(3) and its synthetic analogs inhibit the spontaneous in vitro immunoglobulin production by SLE-derived PBMC." **Clin Immunol** 99(11286544): 82-93.

Liu, M. and Freedman, L. P. (**1994**). "Transcriptional synergism between the vitamin D3 receptor and other nonreceptor transcription factors." **Mol Endocrinol** 8(12): 1593-1604.

Lleo, A., Battezzati, P. M., Selmi, C., Gershwin, M. E. and Podda, M. (**2008**). "Is autoimmunity a matter of sex?" **Autoimmun Rev** 7(8): 626-630.

Llorca, J., Guerrero, P., Prieto-Salceda, D. and Dierssen-Sotos, T. (**2005**). "Mortality of multiple sclerosis in Spain: demonstration of a north-South gradient." **Neuroepidemiology** 24(15650319): 135-140.

Lublin, F. D. and Reingold, S. C. (**1996**). "Defining the clinical course of multiple sclerosis: results of an international survey. National Multiple Sclerosis Society (USA) Advisory Committee on Clinical Trials of New Agents in Multiple Sclerosis." **Neurology** 46(4): 907-911.

Majewski, J. and Pastinen, T. (**2011**). "The study of eQTL variations by RNA-seq: from SNPs to phenotypes." **Trends Genet** 27(2): 72-79.

Martin, G. S., Mannino, D. M., Eaton, S. and Moss, M. (**2003**). "The epidemiology of sepsis in the United States from 1979 through 2000." **N Engl J Med** 348(16): 1546-1554.

Martineau, A. R., Timms, P. M., Bothamley, G. H., Hanifa, Y., Islam, K., Claxton, A. P., Packe, G. E., et al. (**2011**). "High-dose vitamin D(3) during intensive-phase antimicrobial treatment of pulmonary tuberculosis: a double-blind randomised controlled trial." **Lancet** 377(9761): 242-250.

Martineau, A. R., Wilkinson, K. A., Newton, S. M., Floto, R. A., Norman, A. W., Skolimowska, K., Davidson, R. N., et al. (**2007**). "IFN-gamma- and TNF-independent vitamin D-inducible human suppression of mycobacteria: the role of cathelicidin LL-37." **J Immunol** 178(11): 7190-7198.

Mata-Granados, J. M., Vargas-Vasserot, J., Ferreiro-Vera, C., Luque de Castro, M. D., Pavon, R. G. and Quesada Gomez, J. M. (**2010**). "Evaluation of vitamin D endocrine system (VDES) status and response to treatment of patients in intensive care units (ICUs) using an on-line SPE-LC-MS/MS method." **J Steroid Biochem Mol Biol** 121(1-2): 452-455.

Matilainen, J. M., Rasanen, A., Gynther, P. and Vaisanen, S. (**2010**). "The genes encoding cytokines IL-2, IL-10 and IL-12B are primary 1alpha,25(OH)2D3 target genes." **J Steroid Biochem Mol Biol** 121(20236616): 142-145.

Mattner, F., Smiroldo, S., Galbiati, F., Muller, M., Di Lucia, P., Poliani, P. L., Martino, G., et al. (**2000**). "Inhibition of Th1 development and treatment of chronic-relapsing experimental allergic encephalomyelitis by a non-hypercalcemic analogue of 1,25-dihydroxyvitamin D(3)." **Eur J Immunol** 30(10671205): 498-508.

Mawer EB, S. K., Lumb GA, Stanbury SW (**1971**). " The metabolism of isotopically labelled vitamin D in man: the influence of the state of vitamin D nutrition." **Clin Sci** 40: 39–53.

McCombe, P. A., Greer, J. M. and Mackay, I. R. (**2009**). "Sexual dimorphism in autoimmune disease." **Curr Mol Med** 9(9): 1058-1079.

McCulloh, R. and Opal, S. M. (**2013**). "Human recombinant lactoferrin for sepsis: too good to be true?" **Crit Care Med** 41(3): 908-909.

McDonnell, D. P., Mangelsdorf, D. J., Pike, J. W., Haussler, M. R. and O'Malley, B. W. (**1987**). "Molecular cloning of complementary DNA encoding the avian receptor for vitamin D." **Science** 235(4793): 1214-1217.

McDowell, T. Y., Amr, S., Langenberg, P., Royal, W., Bever, C., Culpepper, W. J. and Bradham, D. D. (**2010**). "Time of birth, residential solar radiation and age at onset of multiple sclerosis." **Neuroepidemiology** 34(4): 238-244.

McKenna, N. J. and O'Malley, B. W. (**2002**). "Combinatorial control of gene expression by nuclear receptors and coregulators." **Cell** 108(11909518): 465-474.

McKinney, J. D., Bailey, B. A., Garrett, L. H., Peiris, P., Manning, T. and Peiris, A. N. (**2011**). "Relationship between vitamin D status and ICU outcomes in veterans." **J Am Med Dir Assoc** 12(3): 208-211.

McLeod, J. G., Hammond, S. R. and Hallpike, J. F. (**1994**). "Epidemiology of multiple sclerosis in Australia. With NSW and SA survey results." **Med J Aust** 160(8295576): 117-122.

Meyer, M. B., Goetsch, P. D. and Pike, J. W. (**2012**). "VDR/RXR and TCF4/beta-catenin cistromes in colonic cells of colorectal tumor origin: impact on c-FOS and c-MYC gene expression." **Mol Endocrinol** 26(1): 37-51.

Moller, S., Laigaard, F., Olgaard, K. and Hemmingsen, C. (**2007**). "Effect of 1,25-dihydroxy-vitamin D3 in experimental sepsis." **Int J Med Sci** 4(4): 190-195.

Morgan, J. W., Sliney, D. J., Morgan, D. M. and Maizel, A. L. (**1999**). "Differential regulation of gene transcription in subpopulations of human B lymphocytes by vitamin D3." **Endocrinology** 140(1): 381-391.

Moromizato, T., Litonjua, A. A., Braun, A. B., Gibbons, F. K., Giovannucci, E. and Christopher, K. B. (**2013**). "Association of Low Serum 25-Hydroxyvitamin D Levels and Sepsis in the Critically Ill." **Crit Care Med**.

Movat, H. Z., Cybulsky, M. I., Colditz, I. G., Chan, M. K. and Dinarello, C. A. (**1987**). "Acute inflammation in gram-negative infection: endotoxin, interleukin 1, tumor necrosis factor, and neutrophils." **Fed Proc** 46(1): 97-104.

Mowry, E. M., Waubant, E., McCulloch, C. E., Okuda, D. T., Evangelista, A. A., Lincoln, R. R., Gourraud, P. A., et al. (**2012**). "Vitamin D status predicts new brain magnetic resonance imaging activity in multiple sclerosis." **Ann Neurol** 72(2): 234-240.

Munger, K. L., Levin, L. I., Hollis, B. W., Howard, N. S. and Ascherio, A. (**2006**). "Serum 25-hydroxyvitamin D levels and risk of multiple sclerosis." **JAMA** 296(23): 2832-2838.

Murayama, A., Takeyama, K., Kitanaka, S., Kodera, Y., Kawaguchi, Y., Hosoya, T. and Kato, S. (**1999**). "Positive and negative regulations of the renal 25-hydroxyvitamin D3 1alpha-hydroxylase gene by parathyroid hormone, calcitonin, and 1alpha,25(OH)2D3 in intact animals." **Endocrinology** 140(10218975): 2224-2231.

Naismith, R. T., Piccio, L., Lyons, J. A., Lauber, J., Tutlam, N. T., Parks, B. J., Trinkaus, K., et al. (**2010**). "Rituximab add-on therapy for breakthrough relapsing multiple sclerosis: a 52-week phase II trial." **Neurology** 74(23): 1860-1867.

Neer, R. M. (**1975**). "The evolutionary significance of vitamin D, skin pigment, and ultraviolet light." **Am J Phys Anthropol** 43(3): 409-416.

Nesseler, N., Defontaine, A., Launey, Y., Morcet, J., Malledant, Y. and Seguin, P. (**2013**). "Long-term mortality and quality of life after septic shock: a follow-up observational study." **Intensive Care Med** 39(5): 881-888.

Netea, M. G. and van der Meer, J. W. (**2011**). "Immunodeficiency and genetic defects of pattern-recognition receptors." **N Engl J Med** 364(1): 60-70.

Neviere, R. (**2013**). Pathophysiology of sepsis. G. Finlay. UpToDate. Waltham, MA, Wolters Kluwer Health (Philadelphia, PA). April, 2013.

Neviere, R. (**2013**). Sepsis and the systemic inflammatory response syndrome: Definitions, epidemiology, and prognosis. G. Finlay. UpTodate. Waltham, MA, Wolters Kluwer Health (Philadelphia, PA). April, 2013.

Nielsen, N. M., Westergaard, T., Rostgaard, K., Frisch, M., Hjalgrim, H., Wohlfahrt, J., Koch-Henriksen, N., et al. (**2005**). "Familial risk of multiple sclerosis: a nationwide cohort study." **Am J Epidemiol** 162(8): 774-778.

Nishikawa, J., Kitaura, M., Matsumoto, M., Imagawa, M. and Nishihara, T. (**1994**). "Difference and similarity of DNA sequence recognized by VDR homodimer and VDR/RXR heterodimer." **Nucleic Acids Res** 22(8065900): 2902-2907.

Norman, A. W. (**2006**). "Minireview: vitamin D receptor: new assignments for an already busy receptor." **Endocrinology** 147(12): 5542-5548.

Olek, M. J. (**2013**). Epidemiology and clinical features of multiple sclerosis in adults. J. F. Dashe. UpToDate. Waltham, MA, Wolters Kluwer Health (Philadelphia, PA). April, 2013.

Olek, M. J. (**2013**). Treatment of relapsing-remitting multiple sclerosis in adults. J. F. Dashe. UpToDate. Waltham, MA, Wolters Kluwer Health (Philadelphia, PA). April, 2013.

Olliver, M., Spelmink, L., Hiew, J., Meyer-Hoffert, U., Henriques-Normark, B. and Bergman, P. (**2013**). "Immunomodulatory effects of vitamin D on innate and adaptive immune responses to Streptococcus pneumoniae." **J Infect Dis**.

Opal, S. M., Laterre, P. F., Francois, B., LaRosa, S. P., Angus, D. C., Mira, J. P., Wittebole, X., et al. (**2013**). "Effect of eritoran, an antagonist of MD2-TLR4, on mortality in patients with severe sepsis: the ACCESS randomized trial." **JAMA** 309(11): 1154-1162.

Oren, Y., Shapira, Y., Agmon-Levin, N., Kivity, S., Zafrir, Y., Altman, A., Lerner, A., et al. (**2010**). "Vitamin D insufficiency in a sunny environment: a demographic and seasonal analysis." **Isr Med Assoc J** 12(12): 751-756.

Orlov, I., Rochel, N., Moras, D. and Klaholz, B. P. (**2012**). "Structure of the full human RXR/VDR nuclear receptor heterodimer complex with its DR3 target DNA." **EMBO J** 31(2): 291-300.

Orton, S. M., Morris, A. P., Herrera, B. M., Ramagopalan, S. V., Lincoln, M. R., Chao, M. J., Vieth, R., et al. (**2008**). "Evidence for genetic regulation of vitamin D status in twins with multiple sclerosis." **Am J Clin Nutr** 88(2): 441-447.

Padkin, A., Goldfrad, C., Brady, A. R., Young, D., Black, N. and Rowan, K. (**2003**). "Epidemiology of severe sepsis occurring in the first 24 hrs in intensive care units in England, Wales, and Northern Ireland." **Crit Care Med** 31(9): 2332-2338.

Park, P. J. (**2009**). "ChIP-seq: advantages and challenges of a maturing technology." **Nat Rev Genet** 10(10): 669-680.

Patel, B. M., Chittock, D. R., Russell, J. A. and Walley, K. R. (**2002**). "Beneficial effects of short-term vasopressin infusion during severe septic shock." **Anesthesiology** 96(3): 576-582.

Pender, M. P. (**2009**). "Does Epstein-Barr virus infection in the brain drive the development of multiple sclerosis?" **Brain** 132(Pt 12): 3196-3198.

Penna, G. and Adorini, L. (**2000**). "1 Alpha,25-dihydroxyvitamin D3 inhibits differentiation, maturation, activation, and survival of dendritic cells leading to impaired alloreactive T cell activation." **J Immunol** 164(10679076): 2405-2411.

Penna, G., Roncari, A., Amuchastegui, S., Daniel, K. C., Berti, E., Colonna, M. and Adorini, L. (**2005**). "Expression of the inhibitory receptor ILT3 on dendritic cells is dispensable for induction of CD4+Foxp3+ regulatory T cells by 1,25-dihydroxyvitamin D3." **Blood** 106(16030186): 3490-3497.

Peres Bota, D., Lopes Ferreira, F., Melot, C. and Vincent, J. L. (**2004**). "Body temperature alterations in the critically ill." **Intensive Care Med** 30(5): 811-816.

Perl, T. M., Dvorak, L., Hwang, T. and Wenzel, R. P. (**1995**). "Long-term survival and function after suspected gram-negative sepsis." **JAMA** 274(4): 338-345.

Phillipson, M. and Kubes, P. (**2011**). "The neutrophil in vascular inflammation." **Nat Med** 17(11): 1381-1390.

Pike, J. W. and Meyer, M. B. (**2010**). "The vitamin D receptor: new paradigms for the regulation of gene expression by 1,25-dihydroxyvitamin D(3)." **Endocrinol Metab Clin North Am** 39(20511050): 255-269.

Pinheiro da Silva, F. and Nizet, V. (**2009**). "Cell death during sepsis: integration of disintegration in the inflammatory response to overwhelming infection." **Apoptosis** 14(4): 509-521.

Polman, C. H., Reingold, S. C., Edan, G., Filippi, M., Hartung, H. P., Kappos, L., Lublin, F. D., et al. (**2005**). "Diagnostic criteria for multiple sclerosis: 2005 revisions to the "McDonald Criteria"." **Ann Neurol** 58(6): 840-846.

Purcell, S., Neale, B., Todd-Brown, K., Thomas, L., Ferreira, M. A., Bender, D., Maller, J., et al. (**2007**). "PLINK: a tool set for whole-genome association and population-based linkage analyses." **Am J Hum Genet** 81(3): 559-575.

Quan, H., Li, B., Couris, C. M., Fushimi, K., Graham, P., Hider, P., Januel, J. M., et al. (**2011**). "Updating and validating the Charlson comorbidity index and score for risk adjustment in hospital discharge abstracts using data from 6 countries." **Am J Epidemiol** 173(6): 676-682.

Quesada, J. M., Serrano, I., Borrego, F., Martin, A., Pena, J. and Solana, R. (**1995**). "Calcitriol effect on natural killer cells from hemodialyzed and normal subjects." **Calcif Tissue Int** 56(2): 113-117.

Quraishi, S. A., Litonjua, A. A., Moromizato, T., Gibbons, F. K., Camargo, C. A., Jr., Giovannucci, E. and Christopher, K. B. (**2013**). "Association between

prehospital vitamin D status and hospital-acquired bloodstream infections." **Am J Clin Nutr**.

Rachez, C. and Freedman, L. P. (**2000**). "Mechanisms of gene regulation by vitamin D(3) receptor: a network of coactivator interactions." **Gene** 246(10767523): 9-21.

Rafalski, V. A., Ho, P. P., Brett, J. O., Ucar, D., Dugas, J. C., Pollina, E. A., Chow, L. M., et al. (**2013**). "Expansion of oligodendrocyte progenitor cells following SIRT1 inactivation in the adult brain." **Nat Cell Biol** 15(6): 614-624.

Ramagopalan, S. V., Dobson, R., Meier, U. C. and Giovannoni, G. (**2010**). "Multiple sclerosis: risk factors, prodromes, and potential causal pathways." **Lancet Neurol** 9(7): 727-739.

Ramagopalan, S. V., Dyment, D. A., Cader, M. Z., Morrison, K. M., Disanto, G., Morahan, J. M., Berlanga-Taylor, A. J., et al. (**2011**). "Rare variants in the CYP27B1 gene are associated with multiple sclerosis." **Ann Neurol** 70(6): 881-886.

Ramagopalan, S. V., Heger, A., Berlanga, A. J., Maugeri, N. J., Lincoln, M. R., Burrell, A., Handunnetthi, L., et al. (**2010**). "A ChIP-seq defined genome-wide map of vitamin D receptor binding: associations with disease and evolution." **Genome Res** 20(10): 1352-1360.

Ramagopalan, S. V., Maugeri, N. J., Handunnetthi, L., Lincoln, M. R., Orton, S. M., Dyment, D. A., Deluca, G. C., et al. (**2009**). "Expression of the multiple sclerosis-associated MHC class II Allele HLA-DRB1*1501 is regulated by vitamin D." **PLoS Genet** 5(2): e1000369.

Rangel-Frausto, M. S., Pittet, D., Costigan, M., Hwang, T., Davis, C. S. and Wenzel, R. P. (**1995**). "The natural history of the systemic inflammatory response syndrome (SIRS). A prospective study." **JAMA** 273(2): 117-123.

Ranieri, V. M., Thompson, B. T., Barie, P. S., Dhainaut, J. F., Douglas, I. S., Finfer, S., Gardlund, B., et al. (**2012**). "Drotrecogin alfa (activated) in adults with septic shock." **N Engl J Med** 366(22): 2055-2064.

Ravid, A., Koren, R., Maron, L. and Liberman, U. A. (**1993**). "1,25(OH)2D3 increases cytotoxicity and exocytosis in lymphokine-activated killer cells." **Mol Cell Endocrinol** 96(1-2): 133-139.

Reddy, T. E., Gertz, J., Pauli, F., Kucera, K. S., Varley, K. E., Newberry, K. M., Marinov, G. K., et al. (**2012**). "Effects of sequence variation on differential allelic transcription factor occupancy and gene expression." **Genome Res** 22(5): 860-869.

Reddy, T. E., Pauli, F., Sprouse, R. O., Neff, N. F., Newberry, K. M., Garabedian, M. J. and Myers, R. M. (**2009**). "Genomic determination of the

glucocorticoid response reveals unexpected mechanisms of gene regulation." **Genome Res** 19(12): 2163-2171.

Reid, D., Toole, B. J., Knox, S., Talwar, D., Harten, J., O'Reilly, D. S., Blackwell, S., et al. (**2011**). "The relation between acute changes in the systemic inflammatory response and plasma 25-hydroxyvitamin D concentrations after elective knee arthroplasty." **Am J Clin Nutr** 93(5): 1006-1011.

Rhee, H. S. and Pugh, B. F. (**2011**). "Comprehensive genome-wide protein-DNA interactions detected at single-nucleotide resolution." **Cell** 147(6): 1408-1419.

Riise, T., Nortvedt, M. W. and Ascherio, A. (**2003**). "Smoking is a risk factor for multiple sclerosis." **Neurology** 61(8): 1122-1124.

Risco, A., del Fresno, C., Mambol, A., Alsina-Beauchamp, D., MacKenzie, K. F., Yang, H. T., Barber, D. F., et al. (**2012**). "p38gamma and p38delta kinases regulate the Toll-like receptor 4 (TLR4)-induced cytokine production by controlling ERK1/2 protein kinase pathway activation." **Proc Natl Acad Sci U S A** 109(28): 11200-11205.

Ross, A. C., Manson, J. E., Abrams, S. A., Aloia, J. F., Brannon, P. M., Clinton, S. K., Durazo-Arvizu, R. A., et al. (**2011**). "The 2011 report on dietary reference intakes for calcium and vitamin D from the Institute of Medicine: what clinicians need to know." **J Clin Endocrinol Metab** 96(1): 53-58.

Rozowsky, J., Abyzov, A., Wang, J., Alves, P., Raha, D., Harmanci, A., Leng, J., et al. (**2011**). "AlleleSeq: analysis of allele-specific expression and binding in a network framework." **Mol Syst Biol** 7: 522.

Rudy, J. and Valafar, F. (**2011**). "Empirical comparison of cross-platform normalization methods for gene expression data." **BMC Bioinformatics** 12: 467.

Sadovnick, A. D., Armstrong, H., Rice, G. P., Bulman, D., Hashimoto, L., Paty, D. W., Hashimoto, S. A., et al. (**1993**). "A population-based study of multiple sclerosis in twins: update." **Ann Neurol** 33(3): 281-285.

Sadovnick, A. D., Ebers, G. C., Dyment, D. A. and Risch, N. J. (**1996**). "Evidence for genetic basis of multiple sclerosis. The Canadian Collaborative Study Group." **Lancet** 347(8656905): 1728-1730.

Sandelin, A., Alkema, W., Engstrom, P., Wasserman, W. W. and Lenhard, B. (**2004**). "JASPAR: an open-access database for eukaryotic transcription factor binding profiles." **Nucleic Acids Res** 32(Database issue): D91-94.

Sands, K. E., Bates, D. W., Lanken, P. N., Graman, P. S., Hibberd, P. L., Kahn, K. L., Parsonnet, J., et al. (**1997**). "Epidemiology of sepsis syndrome in 8 academic medical centers." **JAMA** 278(3): 234-240.

Sasse, K. C., Nauenberg, E., Long, A., Anton, B., Tucker, H. J. and Hu, T. W. (**1995**). "Long-term survival after intensive care unit admission with sepsis." **Crit Care Med** 23(6): 1040-1047.

Sawcer, S., Hellenthal, G., Pirinen, M., Spencer, C. C., Patsopoulos, N. A., Moutsianas, L., Dilthey, A., et al. (**2011**). "Genetic risk and a primary role for cell-mediated immune mechanisms in multiple sclerosis." **Nature** 476(7359): 214-219.

Scalfari, A., Neuhaus, A., Degenhardt, A., Rice, G. P., Muraro, P. A., Daumer, M. and Ebers, G. C. (**2010**). "The natural history of multiple sclerosis: a geographically based study 10: relapses and long-term disability." **Brain** 133(Pt 7): 1914-1929.

Schmidt, G. and Mandel, J. (**2013**). Evaluation and management of severe sepsis and septic shock in adults. G. Finlay. Waltham, MA, Wolters Kluwer Health (Philadelphia, PA). April, 2013.

Schrier, R. W. and Wang, W. (**2004**). "Acute renal failure and sepsis." **N Engl J Med** 351(2): 159-169.

Sellebjerg, F., Krakauer, M., Khademi, M., Olsson, T. and Sorensen, P. S. (**2012**). "FOXP3, CBLB and ITCH gene expression and cytotoxic T lymphocyte antigen 4 expression on CD4(+) CD25(high) T cells in multiple sclerosis." **Clin Exp Immunol** 170(2): 149-155.

Semeraro, N., Ammollo, C. T., Semeraro, F. and Colucci, M. (**2012**). "Sepsis, thrombosis and organ dysfunction." **Thromb Res** 129(3): 290-295.

Seok, J., Warren, H. S., Cuenca, A. G., Mindrinos, M. N., Baker, H. V., Xu, W., Richards, D. R., et al. (**2013**). "Genomic responses in mouse models poorly mimic human inflammatory diseases." **Proc Natl Acad Sci U S A** 110(9): 3507-3512.

Shirakawa, A. K., Nagakubo, D., Hieshima, K., Nakayama, T., Jin, Z. and Yoshie, O. (**2008**). "1,25-dihydroxyvitamin D3 induces CCR10 expression in terminally differentiating human B cells." **J Immunol** 180(5): 2786-2795.

Simon, K. C., Munger, K. L. and Ascherio, A. (**2012**). "Vitamin D and multiple sclerosis: epidemiology, immunology, and genetics." **Curr Opin Neurol** 25(3): 246-251.

Simpson, S., Jr., Blizzard, L., Otahal, P., Van der Mei, I. and Taylor, B. (**2011**). "Latitude is significantly associated with the prevalence of multiple sclerosis: a meta-analysis." **J Neurol Neurosurg Psychiatry**.

Singer, M., De Santis, V., Vitale, D. and Jeffcoate, W. (**2004**). "Multiorgan failure is an adaptive, endocrine-mediated, metabolic response to overwhelming systemic inflammation." **Lancet** 364(9433): 545-548.

Smolders, J., Damoiseaux, J., Menheere, P., Tervaert, J. W. and Hupperts, R. (**2009**). "Association study on two vitamin D receptor gene polymorphisms and vitamin D metabolites in multiple sclerosis." **Ann N Y Acad Sci** 1173: 515-520.

Smolders, J., Damoiseaux, J., Menheere, P., Tervaert, J. W. and Hupperts, R. (**2009**). "Fok-I vitamin D receptor gene polymorphism (rs10735810) and vitamin D metabolism in multiple sclerosis." **J Neuroimmunol** 207(1-2): 117-121.

Smolders, J., Peelen, E., Thewissen, M., Cohen Tervaert, J. W., Menheere, P., Hupperts, R. and Damoiseaux, J. (**2010**). "Safety and T cell modulating effects of high dose vitamin D3 supplementation in multiple sclerosis." **PLoS One** 5(12): e15235.

Smolders, J., Peelen, E., Thewissen, M., Menheere, P., Tervaert, J. W., Hupperts, R. and Damoiseaux, J. (**2009**). "The relevance of vitamin D receptor gene polymorphisms for vitamin D research in multiple sclerosis." **Autoimmun Rev** 8(7): 621-626.

Smolders, J., Thewissen, M., Peelen, E., Menheere, P., Tervaert, J. W. C., Damoiseaux, J. and Hupperts, R. (**2009**). "Vitamin D status is positively correlated with regulatory T cell function in patients with multiple sclerosis." **PLoS One** 4(19675671).

Spanier, J. A., Nashold, F. E., Olson, J. K. and Hayes, C. E. (**2012**). "The Ifng gene is essential for Vdr gene expression and vitamin D(3)-mediated reduction of the pathogenic T cell burden in the central nervous system in experimental autoimmune encephalomyelitis, a multiple sclerosis model." **J Immunol** 189(6): 3188-3197.

Spencer, T. E., Jenster, G., Burcin, M. M., Allis, C. D., Zhou, J., Mizzen, C. A., McKenna, N. J., et al. (**1997**). "Steroid receptor coactivator-1 is a histone acetyltransferase." **Nature** 389(9296499): 194-198.

Spielman, R. S. and Ewens, W. J. (**1996**). "The TDT and other family-based tests for linkage disequilibrium and association." **Am J Hum Genet** 59(5): 983-989.

Staples, J., Ponsonby, A. L. and Lim, L. (**2010**). "Low maternal exposure to ultraviolet radiation in pregnancy, month of birth, and risk of multiple sclerosis in offspring: longitudinal analysis." **BMJ** 340: c1640.

Stewart, C. A., Horton, R., Allcock, R. J., Ashurst, J. L., Atrazhev, A. M., Coggill, P., Dunham, I., et al. (**2004**). "Complete MHC haplotype sequencing for common disease gene mapping." **Genome Res** 14(6): 1176-1187.

Stoffels, K., Overbergh, L., Giulietti, A., Verlinden, L., Bouillon, R. and Mathieu, C. (**2006**). "Immune regulation of 25-hydroxyvitamin-D3-1alpha-hydroxylase in human monocytes." **J Bone Miner Res** 21(1): 37-47.

Sundqvist, E., Baarnhielm, M., Alfredsson, L., Hillert, J., Olsson, T. and Kockum, I. (**2010**). "Confirmation of association between multiple sclerosis and CYP27B1." **Eur J Hum Genet** 18(12): 1349-1352.

Tagami, T., Lutz, W. H., Kumar, R. and Jameson, J. L. (**1998**). "The interaction of the vitamin D receptor with nuclear receptor corepressors and coactivators." **Biochem Biophys Res Commun** 253(9878542): 358-363.

Takeda, A., Hamano, S., Yamanaka, A., Hanada, T., Ishibashi, T., Mak, T. W., Yoshimura, A., et al. (**2003**). "Cutting edge: role of IL-27/WSX-1 signaling for induction of T-bet through activation of STAT1 during initial Th1 commitment." **J Immunol** 170(10): 4886-4890.

Taylor, B. V., Lucas, R. M., Dear, K., Kilpatrick, T. J., Pender, M. P., van der Mei, I. A., Chapman, C., et al. (**2010**). "Latitudinal variation in incidence and type of first central nervous system demyelinating events." **Mult Scler** 16(4): 398-405.

Teles, R. M., Graeber, T. G., Krutzik, S. R., Montoya, D., Schenk, M., Lee, D. J., Komisopoulou, E., et al. (**2013**). "Type I interferon suppresses type II interferon-triggered human anti-mycobacterial responses." **Science** 339(6126): 1448-1453.

Templer, D. I., Trent, N. H., Spencer, D. A., Trent, A., Corgiat, M. D., Mortensen, P. B. and Gorton, M. (**1992**). "Season of birth in multiple sclerosis." **Acta Neurol Scand** 85(2): 107-109.

Thomas, M. K., Lloyd-Jones, D. M., Thadhani, R. I., Shaw, A. C., Deraska, D. J., Kitch, B. T., Vamvakas, E. C., et al. (**1998**). "Hypovitaminosis D in medical inpatients." **N Engl J Med** 338(12): 777-783.

Tiosano, D. and Gepstein, V. (**2012**). "Vitamin D action: lessons learned from hereditary 1,25-dihydroxyvitamin-D-resistant rickets patients." **Curr Opin Endocrinol Diabetes Obes** 19(6): 452-459.

Tiosano, D., Wildbaum, G., Gepstein, V., Verbitsky, O., Weisman, Y., Karin, N. and Eztioni, A. (**2013**). "The role of vitamin D receptor in innate and adaptive immunity: a study in hereditary vitamin D-resistant rickets patients." **J Clin Endocrinol Metab** 98(4): 1685-1693.

Toubi, E., Nussbaum, S., Staun-Ram, E., Snir, A., Melamed, D., Hayardeny, L. and Miller, A. (**2012**). "Laquinimod modulates B cells and their regulatory effects on T cells in multiple sclerosis." **J Neuroimmunol** 251(1-2): 45-54.

van Etten, E., Verlinden, L., Giulietti, A., Ramos-Lopez, E., Branisteanu, D. D., Ferreira, G. B., Overbergh, L., et al. (**2007**). "The vitamin D receptor gene FokI polymorphism: functional impact on the immune system." **Eur J Immunol** 37(17274004): 395-405.

van Halteren, A. G. S., Tysma, O. M., van Etten, E., Mathieu, C. and Roep, B. O. (**2004**). "1alpha,25-dihydroxyvitamin D3 or analogue treated dendritic cells modulate human autoreactive T cells via the selective induction of apoptosis." **J Autoimmun** 23(15501394): 233-239.

Vandamme, D., Landuyt, B., Luyten, W. and Schoofs, L. (**2012**). "A comprehensive summary of LL-37, the factotum human cathelicidin peptide." **Cell Immunol** 280(1): 22-35.

Vanhooke, J. L., Benning, M. M., Bauer, C. B., Pike, J. W. and DeLuca, H. F. (**2004**). "Molecular structure of the rat vitamin D receptor ligand binding domain complexed with 2-carbon-substituted vitamin D3 hormone analogues and a LXXLL-containing coactivator peptide." **Biochemistry** 43(15065852): 4101-4110.

Viglietta, V., Baecher-Allan, C., Weiner, H. L. and Hafler, D. A. (**2004**). "Loss of functional suppression by CD4+CD25+ regulatory T cells in patients with multiple sclerosis." **J Exp Med** 199(7): 971-979.

Villagra, A., Cruzat, F., Carvallo, L., Paredes, R., Olate, J., van Wijnen, A. J., Stein, G. S., et al. (**2006**). "Chromatin remodeling and transcriptional activity of the bone-specific osteocalcin gene require CCAAT/enhancer-binding protein beta-dependent recruitment of SWI/SNF activity." **J Biol Chem** 281(16772287): 22695-22706.

Vincent, J. L., Opal, S. M., Marshall, J. C. and Tracey, K. J. (**2013**). "Sepsis definitions: time for change." **Lancet** 381(9868): 774-775.

Vincent, J. L., Zhang, H., Szabo, C. and Preiser, J. C. (**2000**). "Effects of nitric oxide in septic shock." **Am J Respir Crit Care Med** 161(6): 1781-1785.

von Essen, M. R., Kongsbak, M., Schjerling, P., Olgaard, K., Odum, N. and Geisler, C. (**2010**). "Vitamin D controls T cell antigen receptor signaling and activation of human T cells." **Nat Immunol** 11(4): 344-349.

Wang, T. J., Zhang, F., Richards, J. B., Kestenbaum, B., van Meurs, J. B., Berry, D., Kiel, D. P., et al. (**2010**). "Common genetic determinants of vitamin D insufficiency: a genome-wide association study." **Lancet** 376(9736): 180-188.

Wang, T. T., Tavera-Mendoza, L. E., Laperriere, D., Libby, E., MacLeod, N. B., Nagai, Y., Bourdeau, V., et al. (**2005**). "Large-scale in silico and microarray-based identification of direct 1,25-dihydroxyvitamin D3 target genes." **Mol Endocrinol** 19(11): 2685-2695.

Ward, P. A. (**2008**). "Sepsis, apoptosis and complement." **Biochem Pharmacol** 76(11): 1383-1388.

Weber, S. U., Schewe, J. C., Lehmann, L. E., Muller, S., Book, M., Klaschik, S., Hoeft, A., et al. (**2008**). "Induction of Bim and Bid gene expression during accelerated apoptosis in severe sepsis." **Crit Care** 12(5): R128.

Weinshenker, B. G., Bass, B., Rice, G. P., Noseworthy, J., Carriere, W., Baskerville, J. and Ebers, G. C. (**1989**). "The natural history of multiple sclerosis: a geographically based study. I. Clinical course and disability." **Brain** 112 ( Pt 1): 133-146.

Weinshenker, B. G. and Ebers, G. C. (**1987**). "The natural history of multiple sclerosis." **Can J Neurol Sci** 14(3): 255-261.

Whitfield, G. K., Dang, H. T. L., Schluter, S. F., Bernstein, R. M., Bunag, T., Manzon, L. A., Hsieh, G., et al. (**2003**). "Cloning of a functional vitamin D receptor from the lamprey (Petromyzon marinus), an ancient vertebrate lacking a calcified skeleton and teeth." **Endocrinology** 144(12746335): 2704-2716.

Willer, C. J., Dyment, D. A., Sadovnick, A. D., Rothwell, P. M., Murray, T. J. and Ebers, G. C. (**2005**). "Timing of birth and risk of multiple sclerosis: population based study." **BMJ** 330(7483): 120.

Winters, B. D., Eberlein, M., Leung, J., Needham, D. M., Pronovost, P. J. and Sevransky, J. E. (**2010**). "Long-term mortality and quality of life in sepsis: a systematic review." **Crit Care Med** 38(5): 1276-1283.

Wu, W., Beilhartz, G., Roy, Y., Richard, C. L., Curtin, M., Brown, L., Cadieux, D., et al. (**2010**). "Nuclear translocation of the 1,25D3-MARRS (membrane associated rapid response to steroids) receptor protein and NFkappaB in differentiating NB4 leukemia cells." **Exp Cell Res** 316(7): 1101-1108.

Xiao, W., Mindrinos, M. N., Seok, J., Cuschieri, J., Cuenca, A. G., Gao, H., Hayden, D. L., et al. (**2011**). "A genomic storm in critically injured humans." **J Exp Med** 208(13): 2581-2590.

Xu, H., Soruri, A., Gieseler, R. K. and Peters, J. H. (**1993**). "1,25-Dihydroxyvitamin D3 exerts opposing effects to IL-4 on MHC class-II antigen expression, accessory activity, and phagocytosis of human monocytes." **Scand J Immunol** 38(8256111): 535-540.

Xue, M. L., Thakur, A., Willcox, M. D. P., Zhu, H., Lloyd, A. R. and Wakefield, D. (**2003**). "Role and regulation of CXC-chemokines in acute experimental keratitis." **Exp Eye Res** 76(12565810): 221-231.

Yenamandra, S. P., Hellman, U., Kempkes, B., Darekar, S. D., Petermann, S., Sculley, T., Klein, G., et al. (**2010**). "Epstein-Barr virus encoded EBNA-3 binds to vitamin D receptor and blocks activation of its target genes." **Cell Mol Life Sci** 67(24): 4249-4256.

Yu, J., Yu, J., Mani, R. S., Cao, Q., Brenner, C. J., Cao, X., Wang, X., et al. (**2010**). "An integrated network of androgen receptor, polycomb, and TMPRSS2-ERG gene fusions in prostate cancer progression." **Cancer Cell** 17(5): 443-454.

Zahar, J. R., Timsit, J. F., Garrouste-Orgeas, M., Francais, A., Vesin, A., Descorps-Declere, A., Dubois, Y., et al. (**2011**). "Outcomes in severe sepsis and patients with septic shock: pathogen species and infection sites are not associated with mortality." **Crit Care Med** 39(8): 1886-1895.

Zanoni, I., Ostuni, R., Capuano, G., Collini, M., Caccia, M., Ronchi, A. E., Rocchetti, M., et al. (**2009**). "CD14 regulates the dendritic cell life cycle after LPS exposure through NFAT activation." **Nature** 460(7252): 264-268.

Zhang, J., Chalmers, M. J., Stayrook, K. R., Burris, L. L., Wang, Y., Busby, S. A., Pascal, B. D., et al. (**2011**). "DNA binding alters coactivator interaction surfaces of the intact VDR-RXR complex." **Nat Struct Mol Biol** 18(5): 556-563.

Zhang, Q., Raoof, M., Chen, Y., Sumi, Y., Sursal, T., Junger, W., Brohi, K., et al. (**2010**). "Circulating mitochondrial DAMPs cause inflammatory responses to injury." **Nature** 464(7285): 104-107.